

Meller
09/1913763

09/913763

L1 **FILE 'REGISTRY'** ENTERED AT 12:44:39 ON 24 MAY 2002
138 S GGGVIQ/SQSP

L2 **FILE 'CAPLUS'** ENTERED AT 12:45:22 ON 24 MAY 2002
45 S L1
L3 1 S L2 AND (LANTIBIOT? OR SALIVARIC?)

L3 ANSWER 1 OF 1 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:283985 CAPLUS
DOCUMENT NUMBER: 134:316074
TITLE: **Lantibiotic salivaricin B**
production from Streptococcus salivarius
INVENTOR(S): Tagg, John Robert; Dierksen, Karen Patricia;
Upton, Mathew
PATENT ASSIGNEE(S): University of Otago, N. Z.; Blis Technologies
Limited
SOURCE: PCT Int. Appl., 30 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001027143	A1	20010419	WO 2000-NZ197	20001012
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
EP 1169340	A1	20020109	EP 2000-970338	20001012
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO			
NO 2001003905	A	20011010	NO 2001-3905	20010810
PRIORITY APPLN. INFO.:			NZ 1999-500261	A 19991012
			WO 2000-NZ197	W 20001012

AB This invention provides an antibacterial protein, **salivaricin B**. **Salivaricin B** is bacteriocidal with respect to, inter alia, *S. pyogenes* and therefore has numerous therapeutic applications. These applications include, but are not limited to, forming part of therapeutic formulations for use in treating or preventing streptococcal infections of the throat.

IT **335200-05-0**
RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(amino acid sequence; **lantibiotic salivaricin B** prodn. from *Streptococcus salivarius*)

IT **335080-41-6**
RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); OCCU (Occurrence); USES (Uses)

Searcher : Shears 308-4994

09/913763

(amino acid sequence; **lantibiotic salivaricin**
B prodn. from Streptococcus salivarius)

IT 335080-40-5

RL: BOC (Biological occurrence); BSU (Biological study,
unclassified); PRP (Properties); THU (Therapeutic use); BIOL
(Biological study); OCCU (Occurrence); USES (Uses)
(**lantibiotic salivaricin** B prodn. from
Streptococcus salivarius)

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR
THIS RECORD. ALL CITATIONS AVAILABLE IN
THE RE FORMAT

E36 THROUGH E38 ASSIGNED

FILE 'REGISTRY' ENTERED AT 12:58:58 ON 24 MAY 2002

L4 3 S E36-E38

L5 3 S L4 AND L1

L5 ANSWER 1 OF 3 REGISTRY COPYRIGHT 2002 ACS

RN 335200-05-0 REGISTRY

CN L-Serine, L-leucyl-L-threonyl-L-leucyl-L-.alpha.-glutamyl-L-.alpha.-
glutamyl-L-leucyl-L-.alpha.-aspartyl-L-asparaginyl-L-valyl-L-
leucylglycyl-L-alanyl-glycylglycylglycyl-L-valyl-L-isoleucyl-L-
glutaminyl-L-threonyl-L-isoleucyl-L-seryl-L-histidyl-L-.alpha.-
glutamyl-L-cysteinyl-L-arginyl-L-methionyl-L-asparaginyl-L-seryl-L-
tryptophyl-L-glutaminyl-L-phenylalanyl-L-leucyl-L-phenylalanyl-L-
threonyl-L-cysteinyl-L-cysteinyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 7: PN: WO0127143 FIG: 2 claimed protein

CN Salivaricin B (Streptococcus salivarius precursor)

CI MAN

SQL 37

SEQ 1 LTLEELDNVL GAGGGVIQTI SHECRMNSWQ FLFTCCS

=====

HITS AT: 13-18

REFERENCE 1: 134:316074

L5 ANSWER 2 OF 3 REGISTRY COPYRIGHT 2002 ACS

RN 335080-41-6 REGISTRY

CN L-Serine, glycylglycylglycyl-L-valyl-L-isoleucyl-L-glutaminyl-L-
threonyl-L-isoleucyl-L-seryl-L-histidyl-L-.alpha.-glutamyl-L-
cysteinyl-L-arginyl-L-methionyl-L-asparaginyl-L-seryl-L-tryptophyl-L-
glutaminyl-L-phenylalanyl-L-leucyl-L-phenylalanyl-L-threonyl-L-
cysteinyl-L-cysteinyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 3: PN: WO0127143 SEQID: 2 claimed sequence

CN Salivaricin B (Streptococcus salivarius)

SQL 25

SEQ 1 GGGVIQTISH ECRMNSWQFL FTCCS

=====

HITS AT: 1-6

REFERENCE 1: 134:316074

L5 ANSWER 3 OF 3 REGISTRY COPYRIGHT 2002 ACS

09/913763

RN 335080-40-5 REGISTRY
CN L-Glutamine, glycyglycyglycyl-L-valyl-L-isoleucyl- (9CI) (CA
INDEX NAME)

OTHER NAMES:

CN 1: PN: WO0127143 SEQID: 1 claimed sequence

SQL 6

SEQ 1 GGGVIQ
=====

HITS AT: 1-6

REFERENCE 1: 134:316074

=> fil hom

FILE 'HOME' ENTERED AT 12:59:24 ON 24 MAY 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

Run on: May 24, 2002, 09:51:36 ; Search time 51.74 Seconds

(without alignments)
53.669 Million cell updates/sec

Title: US-09-913-763-3

Sequence: 1 GGGVIQTISHECRMNSWQFLFTCCS 25

Scoring table: BLOSUM62

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
Maximum Match	100%	

max limit match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*

1:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1980.DAT *
2:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1981.DAT *
3:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1982.DAT *
4:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1983.DAT *
5:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1984.DAT *
6:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1985.DAT *
7:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1986.DAT *
8:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1987.DAT *
9:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1988.DAT *
10:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1989.DAT *
11:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1990.DAT *
12:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1991.DAT *
13:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1992.DAT *
14:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1993.DAT *
15:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1994.DAT *
16:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1995.DAT *
17:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1996.DAT *
18:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1997.DAT *
19:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1998.DAT *
20:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL1999.DAT *
21:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL2000.DAT *
22:	/SIDSL/gcgcata/hold-genseq/genseqp-emb1/AAL2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	145	100.0	37	22	AAB63356	S. salivarius anti
2	126	86.9	51	19	AAW54813	S. mutans mutacin
3	126	86.9	51	20	AAV03212	Amino acid sequenc
4	115	79.3	25	18	AAW36014	M. varians bacteri
5	115	79.3	47	18	AAW36013	M. varians bacteri
6	115	79.3	47	19	AAW54815	S. mutans mutacin
7	105	72.4	53	19	AAW54812	S. mutans mutacin
8	100	69.0	51	19	AAW54816	S. mutans mutacin
9	100	69.0	51	20	AAV03213	Amino acid sequenc
10	59	40.7	67	18	AAW12808	Hepatitis GB virus
11	59	40.7	120	17	AAK90822	HCV variant EB20-2

12	59	40.7	120	17	AA8B7565	Hepatitis G virus
13	59	40.7	120	19	AA8B0181	Hepatitis G virus
14	59	40.7	120	19	AA876087	HCV variant E820-2-2
15	59	40.7	120	20	AA894584	HCV polypeptide seq
16	59	40.7	120	20	AA892788	US856134 Seq ID 1
17	59	40.7	120	20	AA894547	Hepatitis G virus
18	59	40.7	131	17	AA890821	HCV variant T55806
19	59	40.7	131	17	AA887554	Hepatitis G virus
20	59	40.7	131	19	AA880180	Hepatitis G virus
21	59	40.7	131	19	AA876086	HCV variant T55806
22	59	40.7	131	20	AA894541	Polynucleotide seq
23	59	40.7	131	20	AA892787	US856134 Seq ID 1
24	59	40.7	131	20	AA894546	Hepatitis G virus
25	59	40.7	139	17	AA890820	HCV variant BG34
26	59	40.7	139	17	AA887553	Hepatitis G virus
27	59	40.7	139	19	AA880179	Hepatitis G virus
28	59	40.7	139	19	AA876085	HCV variant BG34 p
29	59	40.7	139	20	AA894540	Polynucleotide seq
30	59	40.7	139	20	AA892786	US856134 Seq ID 1
31	59	40.7	139	20	AA894545	Hepatitis G virus
32	59	40.7	2873	17	AA890766	HCV-PNP 2161 polyp
33	59	40.7	2873	17	AA887559	Hepatitis virus cl
34	59	40.7	2873	18	AA834983	Hepatitis G virus
35	59	40.7	2873	19	AA860148	HCV-PNP2161 varian
36	59	40.7	2873	19	AA876054	HCV isolate PNP 21
37	59	40.7	2873	20	AA894511	Protein encoded by
38	59	40.7	2873	20	AA892755	US856134 Seq ID 1
39	59	40.7	2873	20	AA894542	Hepatitis G virus
40	59	40.7	2910	17	AA890797	HCV-JC variant pol
41	59	40.7	2910	19	AA887566	Hepatitis G virus
42	59	40.7	2910	19	AA860182	Hepatitis G virus
43	59	40.7	2910	19	AA876088	HCV variant JC prot
44	59	40.7	2910	20	AA894559	Polynucleotide seq
45	59	40.7	2910	20	AA892789	US856134 Seq ID 1

ALIGNMENTS

	RESULT	1
ID	AAB62356	
AC	AAB62356 standard; Protein: 37 AA.	
XX		
DT	29-JUN-2001 (first entry)	
XX		
DE	S. salivarius antibacterial protein salivaricin B.	
KW	Salivaricin B; antibacterial; bacteriocidal; Streptococcus; infection;	
KW	upper respiratory tract; medicament; dairy product; milk; confectionery	
KX	bacteriocin-like inhibitory substance; BLIS.	
XX		
OS	Streptococcus salivarius.	
FH	Key	Location/Qualifiers
FT	Peptide	1..12
FT	Protein	/note= "leader sequence"
FT		13..37
FT		/note= "specifically claimed mature protein sequence"
PN	WO200127143-A1.	
PD	19-APR-2001.	
PF	12-OCT-2000; 2000MO-NZ00197.	
PR	12-OCT-1999; 99NZ-0500261.	
XX		
PA	(UYOF-) UNIV OTAGO.	
PA	(BLIS-) BLIS TECHNOLOGIES LTD.	
Tagg JR,	Dierksen KP,	Upton M;
PI		

XX WPI: 2001-281973/29.
 DR N-PSDB; AAF57681.
 XX
 PT Streptococcus salivarius antibacterial protein, known as salivarin B,
 PT useful as an antibiotic for treating infections of the upper
 PT respiratory tract caused by streptococcal organisms, especially
 PT Streptococcus pyogenes -
 XX
 PS Claim 3; Fig 2; 29pp; English.
 XX
 CC The invention provides an antibacterial protein, salivarin B, derived
 CC from Streptococcus salivarius. Salivarin B is bacteriocidal with
 CC respect to Streptococcus pyogenes. Salivarin B is useful for treating
 CC an individual against infections of the upper respiratory tract caused
 CC by streptococcal organisms. An antibacterial composition comprising
 CC salivarin B or an organism which can express salivarin B may be part
 CC of a medicament such as syrup, mouthwash, gargle, toothpaste or mouth
 CC spray. The antibacterial protein or the organism expressing the protein
 CC may be included in a dairy product based food or drink (milk powder, milk
 CC biscuits, milk, yoghurt or cheese, flavored milk) or in confectionery
 CC (e.g. a chewing gum). The antibacterial agents such as bacteriocin-like
 CC one or more secondary antibacterial agents such as bacteriocin-like
 CC inhibitory substance(s) (BLIS). The present sequence represents the
 CC antibacterial protein, salivarin B.
 CC
 XX
 SQ Sequence 37 AA;

Query Match 100.0%; Score 145; DB 22; Length 37;
 Best Local Similarity 100.0%; Pred. No. 2.4e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQTSHECRMSWQFLTCSS 25
 Db 13 gggvltqtshecrmswqfltcscs 37

RESULT 2

AAW54813
 ID AAW54813 standard; Protein; 51 AA.

AC AAW54813;

DT 24-SEP-1998 (first entry)

DE S. mutans mutacin II protein LcndRI.

KW Mutacin; Streptococcus mutans; Gram-positive; dental caries.

OS Streptococcus mutans.

PN WO9817685-A1.

PD 30-APR-1998.

PF 22-OCT-1997; 97WO-US19282.

PR 23-OCT-1996; 96US-0736334.

PA (UABR-) UAB RES FOUND.

PI Caulfield FW, Novak J;

DR WPI: 1998-261424/23.

XX New lantionine-containing antibacterial peptide from Streptococcus
 PT mutans - useful in, e.g. therapeutics for control of Gram-positive
 PT infections, including dental caries
 XX
 PS Disclosure: Page 95; 11pp; English.
 CC The mutacin II protein is one of 5 different proteins which express

CC mutacin isolated from Streptococcus mutans. The mutacin is used
 CC therapeutically to kill Gram-positive bacteria and to treat or inhibit
 CC dental caries, especially when formulated in toothpaste or mouthwash.
 CC The mutacin can be produced in pure form and recombinant methods allow
 CC its overexpression for large scale production.
 CC
 XX
 SQ Sequence 51 AA;

Query Match 86.9%; Score 126; DB 19; Length 51;
 Best Local Similarity 84.0%; Pred. No. 1.5e-10;
 Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVITQTSHECRMSWQFLTCSS 25
 Db 27 gggvltqtshecrmswqfltcscs 51

RESULT 3

AAO3212
 ID AAO3212 standard; Protein; 51 AA.

AC AAO3212;

DT 03-AUG-1999 (first entry)

DE Amino acid sequence of lactacin 481.

KW Sublancin 168; antimicrobial; food preservative; pro-sublancin 168;

KW Gram-positive bacteria; pre-sublancin 168; lactacin 481.

OS unknown.

PN WO9903352-A1.

PD 28-JAN-1999.

PF 17-JUL-1998; 98WO-US14547.

PR 18-JUL-1997; 97US-0053035.

PA (UYMA-) UNITV MARYLAND BALTIMORE.

PI Hansen JN;

DR WPI: 1999-131752/11.

XX New antimicrobial peptide, sublancin 168, from Bacillus subtilis -
 PT used for, e.g. treatment of infections caused by Gram negative
 PT bacteria and as food preservative
 XX
 PS Disclosure: Page 57; 71pp; English.

CC This is the amino acid sequence of lactacin 481 used in the method
 CC of the invention involving the use of prosublancin 168. The peptide
 CC designated sublancin 168, is an antimicrobial useful for treating
 CC infections and preserving food against spoilage bacteria,
 CC particularly Gram-positive bacteria. Pro-sublancin 168 and
 CC pre-sublancin 168, are the precursors of sublancin 168. Sublancin
 CC 168 is very stable at low pH and can be autoclaved without damage.
 CC It does not decompose after 2 years in aqueous solution of about
 CC neutral pH.
 CC
 XX
 SQ Sequence 51 AA;

Query Match 86.9%; Score 126; DB 20; Length 51;
 Best Local Similarity 84.0%; Pred. No. 1.5e-10;
 Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVITQTSHECRMSWQFLTCSS 25
 Db 27 gggvltqtshecrmswqfltcscs 51

```

RESULT 4
ID AAM36014 standard; peptide; 25 AA.
XX
AC AAM36014;
XX
DT 24-FEB-1998 (first entry)
XX
DE M. varians bacteriocin mature peptide sequence.
XX
KM Bacteriocin; Micrococcus varians; probe; PCR; amplification; primer;
KM Lactococcus lactis; lactacin; inhibitor; Listeria monocytogenes; spore;
KM Enterobacteriaceae; Leuconostoc; Streptococcus thermophilus; Bacillus; food;
KM Enterococcus faecalis; Clostridium; gram-negative; contamination;
KM cosmetic; pathogen; skin; dentifrice; Streptococcus sobrinus.
XX
OS Micrococcus varians.
XX
PN EP759469-A1.
XX
PD 26-FEB-1997.
XX
PF 07-AUG-1995; 95EP-0810497.
XX
PR 07-AUG-1995; 95EP-0810497.
XX
PA (NEST ) SOC PROD NESTLE SA.
XX
PI Mollat B, Peel J, Pridmore D, Rekhif N, Suri B;
XX
DR WPI: 1997-147518/14.
XX
DR N-PSDB; AAT94516.
XX
PT Micrococcus varians bacteriocin - useful in mfr. of foods and
XX cosmetics
XX
PS Claim 1; Page 18; 22pp; French.
XX
CC This is the amino acid sequence of the mature peptide from a novel
CC bacteriocin from Micrococcus varians. The protein was isolated from the
CC strains CNCM I-1586 and I-1587 and sequenced. Sequence analysis shows
CC that the protein has a molecular weight of 2659 daltons and the
CC N-terminal 5 amino acids are identical to the lactacin 481 from
CC Lactococcus lactis. The amino acid sequence was used to generate probes
CC and primers used to isolate the sequence encoding the complete protein
CC (AAM36013). The mature protein's activity is not pH dependent, and it
CC is thermostable, retaining its full activity even after incubation at
CC 100 deg. C for 60 mins. The bacteriocin and M. varians strains producing
CC it are useful as inhibitors of a large number of bacterial species,
CC including Listeria monocytogenes, Lactococcus lactis, Lactobacillus
CC spp., Leuconostoc spp., Streptococcus thermophilus and Enterococcus
CC faecalis subsp. faecalis. They are also active against spores and
CC vegetative cells of Bacillus spp. and Clostridium spp., but do not
CC inhibit growth of gram-negative bacteria. The bacteriocin can be used in
CC the manufacture of foods, e.g. salami, cheese and desserts, to control
CC bacterial contamination, and in the manufacture of cosmetics, e.g. to
CC control pathogenic skin bacteria or in dentifrice capable of inhibiting
CC Streptococcus sobrinus.
XX
SQ Sequence 25 AA;

```

Query Match 79.3%; Score 115; DB 18; Length 25;
 Best Local Similarity 80.0%; Pred. No. 2.6e-09;
 Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 GGGVYQTSHECRNMSWQFLTCSS 25
DB 1 ||||| ||||| ||||| ||||| |||||
1 gsgvypitshchmsfgyfctcs 25

```

```

RESULT 5
ID AAM36013 standard; Protein; 47 AA.
XX
AC AAM36013;
XX
DT 24-FEB-1998 (first entry)
XX
DE M. varians bacteriocin protein sequence.
XX
KM Bacteriocin; Micrococcus varians; probe; PCR; amplification; primer;
KM Lactococcus lactis; lactacin; inhibitor; Listeria monocytogenes; spore;
KM Enterobacteriaceae; Leuconostoc; Streptococcus thermophilus; Bacillus; food;
KM Enterococcus faecalis; Clostridium; gram-negative; contamination;
KM cosmetic; pathogen; skin; dentifrice; Streptococcus sobrinus.
XX
OS Micrococcus varians.
XX
FH Key Location/Qualifiers
FT Peptide 1..22
FT /note= "signal peptide"
FT Protein 23..47
FT /note= "mature protein"
XX
PN EP759469-A1.
XX
PD 26-FEB-1997.
XX
PF 07-AUG-1995; 95EP-0810497.
XX
PR 07-AUG-1995; 95EP-0810497.
XX
PA (NEST ) SOC PROD NESTLE SA.
XX
PI Mollat B, Peel J, Pridmore D, Rekhif N, Suri B;
XX
DR WPI: 1997-147518/14.
XX
DR N-PSDB; AAT94516.
XX
PT Micrococcus varians bacteriocin - useful in mfr. of foods and
XX cosmetics
XX
PS Claim 4-6; Page 19; 22pp; French.
XX
CC This is the amino acid sequence of a novel bacteriocin from Micrococcus
CC varians. The protein was isolated from the strains CNCM I-1586 and
CC I-1587 and sequenced. A degenerate oligonucleotide probe (AAT94517) was
CC synthesised and used to probe genomic DNA from strains I-1586 and
CC I-1587. A 1.8 kb BamHI band was detected. A BamHI-restricted genomic DNA
CC library was constructed from which the inserts were amplified with
CC primers AAT94519-20 linked to sequence AAT94518 which corresponds to a
CC fragment of the Lactococcus lactis lactacin 481 gene. The amplification
CC resulted in the isolation of this sequence. The bacteriocin and
CC M. varians strains producing it are useful as inhibitors of a large
CC number of bacterial species, including Listeria monocytogenes,
CC Lactococcus lactis, Lactobacillus spp., Leuconostoc spp., Streptococcus
CC thermophilus and Enterococcus faecalis subsp. faecalis. They are also
CC active against spores and vegetative cells of Bacillus spp. and
CC Clostridium spp., but do not inhibit growth of gram-negative bacteria.
CC The bacteriocin can be used in the manufacture of foods, e.g. salami,
CC cheese and desserts, to control bacterial contamination, and in the
CC manufacture of cosmetics, e.g. to control pathogenic skin bacteria or
CC in dentifrice capable of inhibiting Streptococcus sobrinus.
XX
SQ Sequence 47 AA;

```

Query Match 79.3%; Score 115; DB 18; Length 47;
 Best Local Similarity 80.0%; Pred. No. 4.8e-09;
 Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 GGGVYQTSHECRNMSWQFLTCSS 25
1 ||||| ||||| ||||| ||||| |||||

```

Db 23 gsgvptlshchmnsfqvftccs 47

RESULT 6
AAW54815
ID AAW54815 standard; Protein; 47 AA.
XX
AC AAW54815;
XX
DT 24-SEP-1998 (first entry)
XX
DE S. mutans mutacin II protein VARA.
XX
KW Mutacin; Streptococcus mutans; Gram-positive; dental caries.
XX
OS Streptococcus mutans.
XX
PN MO9817685-A1.
XX
PD 30-APR-1998.
XX
PF 22-OCT-1997; 97WO-US19282.
XX
PR 23-OCT-1996; 96US-0736334.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Caulfield PW, Novak J;
XX
DR WPI; 1998-261424/23.
XX
PT New lantionline-containing antibacterial peptide from Streptococcus
PT mutans - useful in, e.g. therapeutics for control of Gram-positive
PS infections, including dental caries
XX
PS Disclosure; Page 96; 111pp; English.
XX
CC The mutacin II protein is one of 5 different proteins which express
CC mutacin isolated from Streptococcus mutans. The mutacin is used
CC therapeutically to kill Gram-positive bacteria and to treat or inhibit
CC dental caries, especially when formulated in toothpaste or mouthwash.
CC The mutacin can be produced in pure form and recombinant methods allow
CC its overexpression for large scale production.
XX
SQ Sequence 47 AA;

Query Match 79.3%; Score 115; DB 19; Length 47;
Best Local Similarity 80.0%; Pred. No. 4.8e-09;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVITOTISHECRMNSMOFLFTCCS 25
1 ||| ||||| |||:|||||
Db 23 gsgvptlshchmnsfqvftccs 47

RESULT 7
AAW54812
ID AAW54812 standard; Protein; 53 AA.
XX
AC AAW54812;
XX
DT 24-SEP-1998 (first entry)
XX
DE S. mutans mutacin II mut A protein.
XX
KW Mutacin; Streptococcus mutans; Gram-positive; dental caries.
XX
OS Streptococcus mutans.
XX
PN MO9817685-A1.
XX
PD 30-APR-1998.

XX
PF 22-OCT-1997; 97WO-US19282.
XX
PR 23-OCT-1996; 96US-0736334.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Caulfield PW, Novak J;
XX
DR WPI; 1998-261424/23.
XX
N-PSDB; AAV26938.
XX
PT New lantionline-containing antibacterial peptide from Streptococcus
PT mutans - useful in, e.g. therapeutics for control of Gram-positive
PT infections, including dental caries
XX
PS Disclosure; Fig 6; 111pp; English.
XX
CC The mutacin II mut A is a protein encoded by a gene isolated from
CC Streptococcus mutans. The mutacin is used therapeutically to kill
CC Gram-positive bacteria and to treat or inhibit dental caries, especially
CC when formulated in toothpaste or mouthwash. The mutacin can be produced
CC in pure form and recombinant methods allow its overexpression for large
CC scale production.
XX
SQ Sequence 53 AA;

Query Match 72.4%; Score 105; DB 19; Length 53;
Best Local Similarity 72.7%; Pred. No. 1.3e-07;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 GVIOTISHECRMNSMOFLFTCC 24
||:|||||||:||||
Db 32 gvpvtvsecrmnsqwhvftcc 53

RESULT 8
AAW54816
ID AAW54816 standard; Protein; 51 AA.
XX
AC AAW54816;
XX
DT 24-SEP-1998 (first entry)
XX
DE S. mutans mutacin II protein SCNA.
XX
KW Mutacin; Streptococcus mutans; Gram-positive; dental caries.
XX
OS Streptococcus mutans.
XX
PN MO9817685-A1.
XX
PD 30-APR-1998.
XX
PF 22-OCT-1997; 97WO-US19282.
XX
PR 23-OCT-1996; 96US-0736334.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Caulfield PW, Novak J;
XX
DR WPI; 1998-261424/23.
XX
PT New lantionline-containing antibacterial peptide from Streptococcus
PT mutans - useful in, e.g. therapeutics for control of Gram-positive
PT infections, including dental caries
XX
PS Disclosure; Page 96; 111pp; English.
XX
CC The mutacin II protein is one of 5 different proteins which express
CC mutacin isolated from Streptococcus mutans. The mutacin is used

Db 29 gviktishechlnwtlatccs 51

XX

PN W09532292-A2.
 XX 30-NOV-1995.
 PD
 XX
 PF 17-MAY-1995; 95WO-US06266.
 XX
 PR 15-FEB-1995; 95US-0389886.
 PR 20-MAY-1994; 94US-0246985.
 PR 03-AUG-1994; 94US-0285561.
 PR 26-OCT-1994; 94US-0329729.
 PR 23-NOV-1994; 94US-0344271.
 PR 16-DEC-1994; 94US-0357509.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI FRY KE, Kim JP, Linnen JM, Murphy FA, Wages J;
 PI Young LM;
 XX
 DR WPI; 1996-097371/10.
 DR N-PSDB; AAT15715.
 XX
 XX Antigen encoded by reverse frame of positive strand RNA virus
 PT used to determine infection by the virus and to vaccinate against
 PT e.g. hepatitis.
 PS
 PS Example 13; Page 249-50; 303pp; English.
 CC The sequences given in AAR90820-22 are encoded by the sense strand
 CC protein coding regions derived from variant strains of hepatitis G virus
 CC (HGV). The reverse frame corresponding to these sequences may encode
 CC antigenic proteins which may be used in the detection of HGV infection,
 CC and also in the production of vaccines. Three distinct immunogenic
 CC regions have been isolated from three different HGV-epitope libraries
 CC based on HGV serum isolates, JC and PNF 2161. All three epitopic
 CC regions, K1-2-3a, K3 and Y10-13, are encoded by the negative strand of
 CC HGV. The antigenic regions encoded by the negative strand are all
 CC contained within relatively short and separate open reading frames. The
 CC Y10 series of clones was isolated from a cDNA library designated env
 CC which was generated by PCR amplification. Sequence independent single
 CC primer amplified (SISPA) PNF 2161 serum DNA was used as template. The
 CC primers given in AAT15652-55 amplify a fragment of the negative strand of
 CC HGV which contains a 139 amino acid ORF. The ORF has a methionine
 CC present at position 22, where the longest open reading frame is 117 amino
 CC acids. Both the cloned Y10 sequences, Y10-13-1 and Y10-13-2 start
 CC downstream of the Met codon. The primers given in AAT15630-37 amplify
 CC a fragment of HGV containing the C-terminal 31 amino acids of the E2
 CC protein and the N-terminal 166 amino acids of NS2, which corresponds to
 CC the K3 clones. Of the twelve clones isolated, eight contained essentially
 CC the same insert. One was the same as these eight, except for a 3
 CC nucleotide insert. Two of the 12 clones were unique long clones and one
 CC was a unique chimera. Analysis of the position of the K1-2-3a sequence
 CC w.r.t. the sequence of the negative strand of HGV indicated that it
 CC is contained within a 100 amino acid ORF that is located in the
 CC negative strand of the NS3 gene of HGV.
 XX
 SQ Sequence 120 AA;
 QY
 Db Query Match 40.7%; Score 59; DB 17; Length 120;
 Best Local Similarity 37.5%; Pred. No. 0.78;
 Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 2 GGVIGTISHCCRMNSWQFLFTCS 25
 Db 82 gallapathacrangyflncca 105
 RESULT 12
 AAR87565
 ID AAR87565 standard; Peptide; 120 AA.
 XX
 AC AAR87565;
 XX

DT 14-AUG-1996 (first entry)
 XX
 DE Hepatitis G virus clone EB20-2 peptide fragment.
 XX
 KW Hepatitis G virus; HGV; Flaviviridae; viral capsid protein; probe;
 KW viral envelope protein; non-structural protein; RNA helicase; antibody;
 KW chymotrypsin like serine protease; RNA dependent RNA polymerase;
 KW HGV genome.
 OS
 XX Hepatitis G virus clone EB20-2.
 PN W09532291-A2.
 XX
 PD 30-NOV-1995.
 XX
 PF 19-MAY-1995; 95WO-US06169.
 XX
 PR 15-FEB-1995; 95US-0389886.
 PR 20-MAY-1994; 94US-0246985.
 PR 03-AUG-1994; 94US-0285543.
 PR 03-AUG-1994; 94US-0285558.
 PR 26-OCT-1994; 94US-0329729.
 PR 23-NOV-1994; 94US-0344271.
 PR 16-DEC-1994; 94US-0357509.
 PR 03-AUG-1994; 94US-0285561.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI FRY KE, Kim JP, Linnen JM, Wages J, Young LM;
 PI WPI; 1996-049299/05.
 DR N-PSDB; AAT08855.
 XX
 XX Hepatitis G Virus isolates, polypeptide(s) and related nucleic acids
 PT - useful for antibody prodn., vaccines and screening sera
 PT
 PS
 PS Example 6; Page 323; 452pp; English.
 CC AAR87563-R87566 represent fragments of the proteins encoded by hepatitis
 CC G virus (HGV) variants. The DNA encoding these sequences were used as a
 CC comparison to confirm that the full length HGV genomic sequence
 CC represented by AAT08812 is the consensus HGV sequence. AAT08812 was
 CC constructed by annealing the sequences (see AAT08832-T08857) obtained
 CC from a lambda gIII library using HGV specific primers. HGV is a member
 CC of the Flaviviridae family of viruses. These viruses possess a single
 CC positive strand RNA genome. The genome is between 9 and 12 kb in length,
 CC and encodes a single polypeptide of 3000-4000 amino acids long. This
 CC polypeptide is cleaved into approximately 10 proteins, including a viral
 CC capsid protein, at least 1 viral envelope protein, and a minimum of 5
 CC non-structural (NS) proteins. The NS proteins include a chymotrypsin
 CC like serine protease, a RNA helicase (NS3) and a RNA dependent RNA
 CC polymerase (NS5). The full length viral sequence can be used to create
 CC probes for the HGV genome, giving a means of detecting HGV in a sample
 CC (especially a serum). The full length genomic sequence can be cloned
 CC into a vector, which can be used to express the encoded protein in a
 CC host cell for isolation and purification. The protein, can then be used
 CC to create antibodies specifically reactive with HGV or a HGV epitope.
 CC The antibodies can be used to detect the presence of HGV in a sample, as
 CC in vaccines. The purified protein sequence can also be used
 XX
 SQ Sequence 120 AA;
 QY
 Db Query Match 40.7%; Score 59; DB 17; Length 120;
 Best Local Similarity 37.5%; Pred. No. 0.78;
 Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 2 GGVIGTISHCCRMNSWQFLFTCS 25
 Db 82 gallapathacrangyflncca 105


```

RESULT 13
AAW80181
ID AAW80181 standard; Protein; 120 AA.
XX
AC AAW80181;
XX
DT 23-DEC-1998 (first entry)
XX
DE Hepatitis G virus (HGV) variant EB20-2.
XX
KW PNF2161 sera; antigen: immunoreaction screening;
KW Non-A Non-B Non-C Non-D Non-E Hepatitis virus; HGV;
KW anti-HGV antibody; vaccine.
XX
OS Hepatitis G virus.
XX
PN US5824507-A.
XX
PD 20-OCT-1998.
XX
PF 19-MAY-1995; 95US-0444733.
XX
PR 19-MAY-1995; 95US-0444733.
PR 20-MAY-1994; 94US-0246985.
PR 03-AUG-1994; 94US-0285543.
PR 03-AUG-1994; 94US-0285561.
PR 26-OCT-1994; 94US-0329729.
PR 23-NOV-1994; 94US-0344271.
PR 16-DEC-1994; 94US-0357509.
PR 15-FEB-1995; 95US-0389886.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Fry KE, Kim JP, Linnen JM, Wages J, Young LM;
XX
DR WPI; 1998-582553/49.
XX
DR N-PSDB; AAV66228.
XX
PT Polypeptide antigens hepatitis G virus - useful as vaccines against
PT the virus and raising antibodies for use in immuno-testing for the
PT virus
XX
PS Example 18; Columns 265-266; 206pp; English.
XX
CC The present sequence represents a Hepatitis G virus (HGV) variant EB20-2.
CC The specification describes HGV antigen clones isolated from PNF2161
CC sera. This sera is obtained from a patient suffering from Non-A Non-B
CC HGV. Non-D Non-E Hepatitis Virus, also known as Hepatitis G virus (HGV).
CC HGV antigens can be used in immunoreaction screening for presence of
CC anti-HGV specific antibodies, especially in patient sera, e.g. enzyme
CC linked immunoassay (ELISA). They can also be used as vaccines, and to
CC raise the polyclonal and monoclonal antibodies.
XX
SQ Sequence 120 AA;
XX
Query Match 40.7%; Score 59; DB 19; Length 120;
Best Local Similarity 37.5%; Pred. No. 0.78; 11; Indels 0; Gaps 0;
Matches 9; Conservative 4; Mismatches 11;
QY 2 GGVIOITISHCCRMNSWQFLFTCCS 25
Db 82 gallapathacrangyfltncca 105

```

```

DE HGV variant EB20-2 protein fragment.
XX
KW Immunoreactive; detection; antibody; alanine aminotransferase;
KW Non-A, Non-B, Non-C, Non-D, Non-E hepatitis virus.
XX
OS Hepatitis G virus.
XX
PN US5766840-A.
XX
PD 16-JUN-1998.
XX
PF 05-JUN-1995; 95US-0466033.
XX
PR 19-MAY-1995; 95US-0444733.
PR 20-MAY-1994; 94US-0246985.
PR 03-AUG-1994; 94US-0285543.
PR 03-AUG-1994; 94US-0285558.
PR 03-AUG-1994; 94US-0285561.
PR 26-OCT-1994; 94US-0329729.
PR 23-NOV-1994; 94US-0344271.
PR 16-DEC-1994; 94US-0357509.
PR 15-FEB-1995; 95US-0389886.
PR 05-JUN-1995; 95US-0466033.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Fry KE, Kim JP, Linnen JM, Wages J, Young LM;
XX
DR WPI; 1998-361677/31.
XX
PT Antibody to hepatitis G virus protein - useful for diagnosis of
PT hepatitis G virus infection
XX
PS Example 6; Column 265-266; 204pp; English.
XX
CC This sequence is a used in a method which detects Non-A, Non-B, Non-C,
CC Non-D, Non-E hepatitis G virus (HGV) in a biological fluid sample, e.g.
CC from a test subject, especially using a kit comprising an anti-HGV
CC antibody preparation and a reporter for detecting binding of an HGV
CC polypeptide antigen to the antibody, preferably where the antibody is
CC monoclonal and/or is attached to a solid support and where the reporter
CC is a labelled monoclonal antibody or a labelled competing antigen. The
CC HGV is characterised by producing elevated serum alanine aminotransferase
CC levels in an infected primate, being serologically distinct from
CC hepatitis A, B, C, D and E viruses and having a viral genome containing a
CC nucleic acid region that is selectively hybridisable with a polypeptide
CC having 203 base pairs as given in the specification.
XX
SQ Sequence 120 AA;
XX
Query Match 40.7%; Score 59; DB 19; Length 120;
Best Local Similarity 37.5%; Pred. No. 0.78; 11; Indels 0; Gaps 0;
Matches 9; Conservative 4; Mismatches 11;
QY 2 GGVIOITISHCCRMNSWQFLFTCCS 25
Db 82 gallapathacrangyfltncca 105

```

```

RESULT 14
AAW76087
ID AAW76087 standard; Protein; 120 AA.
XX
AC AAW76087;
XX
DT 06-NOV-1998 (first entry)
XX

```

```

RESULT 15
AAW94514
ID AAW94514 standard; Protein; 120 AA.
XX
AC AAW94514;
XX
DT 22-JUL-1999 (first entry)
XX
DE HGV polypeptide sequence.
KW Hepatitis G virus; HGV; HGV-E2 antigen; vaccine; viral replication;
KW treatment; HGV infection; antisense.
XX

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:52:04 ; Search time 21.74 Seconds
(without alignments)
28.088 Million cell updates/sec

Title: US-09-913-763-3
Perfect score: 145
Sequence: 1 GGGVITQISHCCRNNSWQFLPTCCS 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCRNUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	86.9	51	2	US-08-736-334B-8
2	115	79.3	25	1	US-08-861-775-1
3	115	79.3	25	1	US-08-833-887-1
4	115	79.3	25	2	US-08-823-603-1
5	115	79.3	25	4	US-09-257-710-1
6	115	79.3	47	1	US-08-861-775-3
7	115	79.3	47	1	US-08-833-887-3
8	115	79.3	47	2	US-08-736-334B-10
9	115	79.3	47	2	US-08-823-603-3
10	115	79.3	47	4	US-09-257-710-3
11	105	72.4	53	2	US-08-736-334B-7
12	100	69.0	51	2	US-08-736-334B-11
13	59	40.7	67	2	US-08-639-857-31
14	59	40.7	120	1	US-08-466-033-181
15	59	40.7	120	1	US-08-444-733-181
16	59	40.7	120	2	US-08-464-134-181
17	59	40.7	120	2	US-08-461-361-181
18	59	40.7	120	2	US-08-485-910-181
19	59	40.7	120	5	PCT-US95-06266-155
20	59	40.7	131	1	US-08-466-033-179
21	59	40.7	131	2	US-08-444-733-179
22	59	40.7	131	2	US-08-464-134-179
23	59	40.7	131	2	US-08-461-361-179
24	59	40.7	131	2	US-08-485-910-179
25	59	40.7	131	5	PCT-US95-06266-153
26	59	40.7	139	1	US-08-466-033-177
27	59	40.7	139	2	US-08-444-733-177

28	59	40.7	139	2	US-08-464-134-177	Sequence 177, App
29	59	40.7	139	2	US-08-461-361-177	Sequence 177, App
30	59	40.7	139	2	US-08-485-910-177	Sequence 177, App
31	59	40.7	139	5	PCT-US95-06266-151	Sequence 151, App
32	59	40.7	2873	1	US-08-466-033-15	Sequence 15, App
33	59	40.7	2873	1	US-08-638-911A-2	Sequence 2, Appl
34	59	40.7	2873	2	US-08-444-733-15	Sequence 15, Appl
35	59	40.7	2873	2	US-08-464-134-15	Sequence 15, Appl
36	59	40.7	2873	2	US-08-461-361-15	Sequence 15, Appl
37	59	40.7	2873	2	US-08-485-910-15	Sequence 15, Appl
38	59	40.7	2873	5	PCT-US95-06266-15	Sequence 15, Appl
39	59	40.7	2910	1	US-08-466-033-183	Sequence 183, App
40	59	40.7	2910	2	US-08-444-733-183	Sequence 183, App
41	59	40.7	2910	2	US-08-464-134-183	Sequence 183, App
42	59	40.7	2910	2	US-08-461-361-183	Sequence 183, App
43	59	40.7	2910	2	US-08-485-910-183	Sequence 183, App
44	59	40.7	2910	5	PCT-US95-06266-157	Sequence 157, App
45	59	40.7	31.7	2	US-08-639-857-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-08-736-334B-8
Sequence 8, Application US/08736334B
Patent No. 5872001
GENERAL INFORMATION:
APPLICANT: Caulfield et al.
TITLE OF INVENTION: NOVEL LANTHONINE ANTIBIOTIC
TITLE OF INVENTION: COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. Benjamin A. Adler
STREET: 8011 Candle Lane
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/736,334B
FILING DATE: October 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Dr. Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5911CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acid residues
TYPE: amino acids
TOPOLOGY: Linear
US-08-736-334B-8

Query Match 86.9%; Score 126; DB 2; Length 51;
Best Local Similarity 84.0%; Pred. No. 6.8e-12;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVITQISHCCRNNSWQFLPTCCS 25
| ||| ||||| ||||| |||||
Db 27 GGGVITQISHCCRNNSWQFLPTCCS 51

RESULT 2

```
US-08-861-775-1
; Sequence 1, Application US/08861775
; Patent No. 575665
; GENERAL INFORMATION:
; APPLICANT: MOLETT, BEAT
; APPLICANT: PEEL, JOHN
; APPLICANT: PRIMORE, DAVID
; APPLICANT: REKIF, NADJI
; APPLICANT: SURI, BRUNO
; TITLE OF INVENTION: BACTERIOCIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: VOGT & O'DONNELL
; STREET: 707 WESTCHESTER AVENUE
; CITY: WHITE PLAINS, NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/861,775
; FILING DATE: 22-MAY-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/693,353
; FILING DATE: 23-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: HARACZ, STEPHEN M.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)328 0055
; TELEFAX: (914)328 0060
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: MICROCOCUS VARIANS
; INDIVIDUAL ISOLATE: TWO CLONES CNCM I-1586 and CNCM
; INDIVIDUAL ISOLATE: I-1587
US-08-861-775-1

Query Match          79.3%; Score 115; DB 1; Length 25;
Best Local Similarity 80.0%; Pred. No. 1.3e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVIOFISHCHRMNSMOFLTCSS 25
DB 1 GSGVIFPISHCHRMNSFOVFTCSS 25

RESULT 3
US-08-833-887-1
; Sequence 1, Application US/08833887
; Patent No. 5763247
; GENERAL INFORMATION:
; APPLICANT: MOLETT, BEAT
; APPLICANT: PEEL, JOHN
; APPLICANT: PRIMORE, DAVID
; APPLICANT: REKIF, NADJI
; APPLICANT: SURI, BRUNO
; TITLE OF INVENTION: BACTERIOCIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: VOGT & O'DONNELL
; STREET: 707 WESTCHESTER AVENUE
```

```
; CITY: WHITE PLAINS, NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,887
; FILING DATE: 10-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HARACZ, STEPHEN M.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)328 0055
; TELEFAX: (914)328 0060
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: MICROCOCUS VARIANS
; INDIVIDUAL ISOLATE: TWO CLONES CNCM I-1586 and CNCM
; INDIVIDUAL ISOLATE: I-1587
US-08-833-887-1

Query Match          79.3%; Score 115; DB 1; Length 25;
Best Local Similarity 80.0%; Pred. No. 1.3e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVIOFISHCHRMNSMOFLTCSS 25
DB 1 GSGVIFPISHCHRMNSFOVFTCSS 25

RESULT 4
US-08-823-603-1
; Sequence 1, Application US/08823603
; Patent No. 5981261
; GENERAL INFORMATION:
; APPLICANT: MOLETT, BEAT
; APPLICANT: PEEL, JOHN
; APPLICANT: PRIMORE, DAVID
; APPLICANT: REKIF, NADJI
; APPLICANT: SURI, BRUNO
; TITLE OF INVENTION: BACTERIOCIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: VOGT & O'DONNELL
; STREET: 707 WESTCHESTER AVENUE
; CITY: WHITE PLAINS, NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,603
; FILING DATE:
; CLASSIFICATION:
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/693,353
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HARACZ, STEPHEN M.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914)328 0055
TELEFAX: (914)328 0060
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: MICROCOCUS VARIANS
INDIVIDUAL ISOLATE: TWO CLONES CNCM I-1586 and CNCM
INDIVIDUAL ISOLATE: I-1587
US-08-823-603-1

Query Match 79.3%; Score 115; DB 2; Length 25;
Best local Similarity 80.0%; Pred. No. 1.3e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVIOITISHECRNMSWQFLFTCCS 25
1 ||| ||||| |||:|||||
DB 1 GSGVPTISHECHMNSFOFVFTCCS 25

RESULT 5
US-09-257-710-1
Sequence 1, Application US/09257710
Patent No. 6150139
GENERAL INFORMATION:
APPLICANT: MOLLET, BEAT
APPLICANT: PEEL, JOHN
APPLICANT: PRIMMORE, DAVID
APPLICANT: REKHIF, NADJI
APPLICANT: SURI, BRUNO
TITLE OF INVENTION: BACTERIOCIN
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: VOGT & O'DONNELL
STREET: 707 WESTCHESTER AVENUE
CITY: WHITE PLAINS, NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10604
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/257,710
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/693,353
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HARACZ, STEPHEN M.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914)328 0055
TELEFAX: (914)328 0060
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: MICROCOCUS VARIANS
INDIVIDUAL ISOLATE: TWO CLONES CNCM I-1586 and CNCM
INDIVIDUAL ISOLATE: I-1587
US-09-257-710-1

Query Match 79.3%; Score 115; DB 4; Length 25;
Best local Similarity 80.0%; Pred. No. 1.3e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVIOITISHECRNMSWQFLFTCCS 25
1 ||| ||||| |||:|||||
DB 1 GSGVPTISHECHMNSFOFVFTCCS 25

RESULT 6
US-08-861-775-3
Sequence 3, Application US/08861775
Patent No. 5756665
GENERAL INFORMATION:
APPLICANT: MOLLET, BEAT
APPLICANT: PEEL, JOHN
APPLICANT: PRIMMORE, DAVID
APPLICANT: REKHIF, NADJI
APPLICANT: SURI, BRUNO
TITLE OF INVENTION: BACTERIOCIN
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: VOGT & O'DONNELL
STREET: 707 WESTCHESTER AVENUE
CITY: WHITE PLAINS, NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10604
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/861,775
FILING DATE: 22-MAY-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/693,353
FILING DATE: 23-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: HARACZ, STEPHEN M.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914)328 0055
TELEFAX: (914)328 0060
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-861-775-3

Query Match 79.3%; Score 115; DB 1; Length 47;
Best local Similarity 80.0%; Pred. No. 2.6e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVIOITISHECRNMSWQFLFTCCS 25
1 ||| ||||| |||:|||||
DB 23 GSGVPTISHECHMNSFOFVFTCCS 47

RESULT 7
US-08-833-887-3

```
; Sequence 3, Application US/08833887
; Patent No. 5763247
; GENERAL INFORMATION:
; APPLICANT: MOLLET, BEAT
; APPLICANT: PEEL, JOHN
; APPLICANT: PRIDMORE, DAVID
; APPLICANT: REKHIF, NADJI
; APPLICANT: SURI, BRUNO
; TITLE OF INVENTION: BACTERIOCIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: VOGT & O'DONNELL
; STREET: 707 WESTCHESTER AVENUE
; CITY: WHITE PLAINS, NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,887
; FILING DATE: 10-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/693,353
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HARACZ, STEPHEN M.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)328 0055
; TELEFAX: (914)328 0060
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-887-3

Query Match          79.3%; Score 115; DB 1; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.6e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVITQTSHECRMNSMQLFTCCS 25
DB 23 GSGVITPTISHCHMNSFQVFTCCS 47

RESULT 8
US-08-736-334B-10
; Sequence 10, Application US/08736334B
; Patent No. 5872601
; GENERAL INFORMATION:
; APPLICANT: Caufield et al.
; TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
```

```
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/736,334B
; FILING DATE: October 23, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Dr. Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5911CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acid residues
; TYPE: amino acids
; TOPOLOGY: linear
; US-08-736-334B-10
```

```
Query Match          79.3%; Score 115; DB 2; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.6e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 GGGVITQTSHECRMNSMQLFTCCS 25
DB 23 GSGVITPTISHCHMNSFQVFTCCS 47
```

```
RESULT 9
US-08-823-603-3
; Sequence 3, Application US/08823603
; Patent No. 5981261
; GENERAL INFORMATION:
; APPLICANT: MOLLET, BEAT
; APPLICANT: PRIDMORE, DAVID
; APPLICANT: REKHIF, NADJI
; APPLICANT: SURI, BRUNO
; TITLE OF INVENTION: BACTERIOCIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: VOGT & O'DONNELL
; STREET: 707 WESTCHESTER AVENUE
; CITY: WHITE PLAINS, NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,603
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/693,353
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HARACZ, STEPHEN M.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)328 0055
; TELEFAX: (914)328 0060
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-823-603-3
```

Query Match 79.3%; Score 115; DB 2; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.6e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVIGTISHRCNMSWQFLFTCCS 25
DB 23 GSGVIPTISHCHMNSFQVFTCCS 47

RESULT 10
US-09-257-710-3
Sequence 3, Application US/09257710
Patent No. 6150139
GENERAL INFORMATION:
APPLICANT: MOLETT, BEAT
APPLICANT: PEEL, JOHN
APPLICANT: PRIDMORE, DAVID
APPLICANT: REKHIF, NADUT
APPLICANT: SURI, BRUNO
TITLE OF INVENTION: BACTERIOCIN
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: VOGT & O'DONNELL
STREET: 707 WESTCHESTER AVENUE
CITY: WHITE PLAINS, NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/257,710
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/693,353
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HARACZ, STEPHEN M.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914)328 0055
TELEFAX: (914)328 0060
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-257-710-3

Query Match 79.3%; Score 115; DB 4; Length 47;
Best Local Similarity 80.0%; Pred. No. 2.6e-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVIGTISHRCNMSWQFLFTCCS 25
DB 23 GSGVIPTISHCHMNSFQVFTCCS 47

RESULT 11
US-08-736-334B-7
Sequence 7, Application US/08736334B
Patent No. 5872001
GENERAL INFORMATION:
APPLICANT: Caufield et al.
TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC
INFORMATION FOR SEQ ID NO: 13

CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. Benjamin A. Adler
STREET: 8011 Candle Lane
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/736,334B
FILING DATE: October 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Dr. Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5911CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acid residues
TYPE: amino acids
TOPOLOGY: linear
US-08-736-334B-7

Query Match 72.4%; Score 105; DB 2; Length 53;
Best Local Similarity 72.7%; Pred. No. 8.5e-09;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 GVIOTISHRCNMSWQFLFTCC 24
DB 32 GVPVTSYECRMSWQHFTCC 53

RESULT 12
US-08-736-334B-11
Sequence 11, Application US/08736334B
Patent No. 5872001
GENERAL INFORMATION:
APPLICANT: Caufield et al.
TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC
INFORMATION FOR SEQ ID NO: 13
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. Benjamin A. Adler
STREET: 8011 Candle Lane
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/736,334B
FILING DATE: October 23, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Dr. Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5911CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acid residues
TYPE: amino acids
TOPOLOGY: linear
US-08-736-334B-11

Query Match 69.0%; Score 100; DB 2; Length 51;
Best Local Similarity 69.6%; Pred. No. 4.4e-08;
Matches 16; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 GGVIQTISHCRNMSWQFLFTCCS 25
||:|||||:|:| |||||
DB 29 GVEKTISSHECHNTWAFATCCS 51

RESULT 13
US-08-639-857-31
Sequence 31, Application US/08639857
Patent No. 5955318

GENERAL INFORMATION:
APPLICANT: Simons, J. N.
APPLICANT: Desai, S. M.
APPLICANT: Mushahwar, I. K.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR CONTROLLING THE
NUMBER OF SEQUENCES: 32
TRANSLATION OF HEPATITIS GB PROTEINS
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Rd
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,857
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Potembksi, Priscilla E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5793.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-0378
TELEFAX: 708-938-2623

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 67 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-639-857-31

Query Match 40.7%; Score 59; DB 2; Length 67;
Best Local Similarity 37.5%; Pred. No. 0.062;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 2 GGVIQTISHCRNMSWQFLFTCCS 25
|:::| || | |||||
DB 44 GALLAPATHACRANGQYFLTNCCA 67

RESULT 14
US-08-466-033-181
Sequence 181, Application US/08466033
Patent No. 5766840

GENERAL INFORMATION:
APPLICANT: Kim, Jungsub P.
APPLICANT: Wages, John
APPLICANT: Young, Lavonne M.
APPLICANT: Fry, Kirk E.
APPLICANT: Linnen, Jeffrey M.
TITLE OF INVENTION: Hepatitis G Virus and Molecular
NUMBER OF SEQUENCES: 277
TRANSLATION OF HEPATITIS GB PROTEINS
CORRESPONDENCE ADDRESS:
ADDRESSEE: Delinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,033
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/389,886
FILING DATE: 15-FEB-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/357,509
FILING DATE: 16-DEC-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,729
FILING DATE: 26-OCT-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,271
FILING DATE: 23-NOV-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/285,558
FILING DATE: 03-AUG-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/285,543
FILING DATE: 03-AUG-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/246,985
FILING DATE: 20-MAY-1994

ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 181:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-033-181

Query Match 40.7%; Score 59; DB 1; Length 120;
Best Local Similarity 37.5%; Pred. No. 0.11;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 2 GGVIQTISHCRNMSWQFLFTCCS 25
|:::| || | |||||
DB 82 GALLAPATHACRANGQYFLTNCCA 105

RESULT 15
US-08-444-733-181


```
Search completed: May 24, 2002, 09:52:05
Job time: 150 sec
```

```

? Sequence 181, Application US/08444733
? Patent No. 5824507
?
? GENERAL INFORMATION:
? APPLICANT: Kim, Jungsuh P.
? APPLICANT: Wages, John
? APPLICANT: Young, Lavonne M.
? APPLICANT: Fry, Kirk E.
? APPLICANT: Linnen, Jeffrey M.
? TITLE OF INVENTION: Hepatitis G Virus and Molecular
? TITLE OF INVENTION: Cloning Thereof
? NUMBER OF SEQUENCES: 277
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Dehlinger & Associates
? STREET: 350 Cambridge Ave., Suite 250
? CITY: Palo Alto
? STATE: CA
? COUNTRY: USA
? ZIP: 94306
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/444,733
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/389,886
? FILING DATE: 15-FEB-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/357,509
? FILING DATE: 16-DEC-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/329,729
? FILING DATE: 26-OCT-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/344,271
? FILING DATE: 23-NOV-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/285,558
? FILING DATE: 03-AUG-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/285,543
? FILING DATE: 03-AUG-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/246,985
? FILING DATE: 20-MAY-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Fabian, Gary R.
? REGISTRATION NUMBER: 33,875
? REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 324-0880
? TELEFAX: (415) 324-0960
? INFORMATION FOR SEQ ID NO: 181:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 120 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
? US-08-444-733-181

```

Query Match	40.7%;	Score 59;	DB 2;	Length 120;
Best Local Similarity	37.5%;	Pred. No. 0.11;		
Matches	9;	Conservative	4;	Mismatches 11; Indels 0; Gaps 0;

```
QY      2 GGVITISHECRMNSWQFLFTCCS 25
          |:::| | | | | | | | | |
Db      82 GAILPATHTACRANGQYFLTNCCA 105
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:52:38 ; Search time 26.69 Seconds

(without alignments)
90.005 Million cell updates/sec

Title: US-09-913-763-3

Perfect score: 145
Sequence: 1 GGGVIGTISHECRMNSMOWFTCCS 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.71:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	126	86.9	51	2	A47342
2	115	79.3	47	2	A58711
3	105	72.4	53	2	J06526
4	100	69.0	51	2	T09004
5	75	51.7	51	2	T09005
6	65	44.8	2970	2	T08839
7	55	37.9	3005	2	T08841
8	52.5	36.2	553	1	TLBPF7
9	51	35.2	412	2	G98217
10	51	35.2	412	2	AD3069
11	51	35.2	413	2	E96035
12	51	35.2	413	2	AF3450
13	50	34.5	1650	2	S53457
14	49.5	34.1	606	2	T47690
15	49	33.8	2358	2	T39569
16	49	33.8	2371	2	T43432
17	47	32.4	337	2	B71340
18	46.5	32.1	311	2	C71612
19	46	31.7	747	2	A84047
20	46	31.7	747	2	T33488
21	46	31.7	864	2	C82019
22	46	31.7	864	2	A81000
23	46	31.7	1064	2	F86182
24	45.5	31.4	85	4	I70119
25	45.5	31.4	1152	2	T31911
26	45	31.0	120	2	D86435
27	45	31.0	295	2	AD2211
28	45	31.0	336	2	G82228
29	45	31.0	352	1	AUKXQ

ALIGNMENTS

```

30      45      31.0      461 2  G64206      hexosephosphate tr
31      45      31.0      634 2  S73461      transport ATP-bind
32      45      31.0      898 2  A40114      fasciclin II precu
33      44.5      30.7      153 2  T46110      hypothetical prote
34      44.5      30.7      186 2  A45189      chitin synthase (E
35      44.5      30.7      486 2  T45763      hypothetical prote
36      44.5      30.7      1308 2  T15280      hypothetical prote
37      44      30.3      60 2  C82533      hypothetical prote
38      44      30.3      573 2  T19880      hypothetical prote
39      44      30.3      650 2  T23175      hypothetical prote
40      44      30.3      898 2  T42131      probable toxR-regu
41      44      30.3      918 2  G86502      S/T protein kinase
42      44      30.3      918 2  C72120      s/t protein kinase
43      43.5      30.0      171 2  J00699      hypothetical 19k p
44      43      29.7      308 1  G64906      glutaminase homolo
45      43      29.7      308 2  C90895      probable glutamina

RESULT 1
A47342
lactacin 481 precursor - Lactococcus lactis subsp. lactis
C:Species: Lactococcus lactis subsp. lactis
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 24-Sep-1999
C:Accession: A47342; S36019
R:Plard, J.C.; Kuipers, O.P.; Rollet, H.S.; Desmazeaud, M.J.; de Vos, W.M.
J. Biol. Chem. 268, 16361-16368, 1993
A:Title: Structure, organization, and expression of the lct gene for lactacin 481, a
A:Reference number: A47342; MUID:93346379
A:Accession: A47342
A:Molecule type: DNA
A:Residues: 1-51 <PIA>
A:Cross-references: EMBL:X71410; NID:g296368; PIDN:CA50534.1; PID:g296369
A>Note: Submitted to the EMBL Data Library, April 1993
R:van den Hooven, H.W.; Lagerweij, F.M.; Heerma, W.; Haverkamp, J.; Plard, J.C.; Hilb
FBS Lett. 391, 317-322, 1996
A:Title: The structure of the lantibiotic lactacin 481 produced by Lactococcus lactis
A:Reference number: A58711; MUID:96350545
A:Contents: annotation; conformation and sulfide bond assignments by (1)H-NMR, mass-s
C:Genetics:
A:Gene: lct
C:Superfamily: unassigned lantibiotic-containing peptides
C:Keywords: antibiolic; lantibiotic
F:1-24/Domain: propeptide #status predicted <PRO>
F:25-51/Product: lactacin 481 #status experimental <MAP>
F:33-38/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status experimental
F:35-49/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F:42-50/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F:48/Modified site: (Z)-dehydrobutyryl (Thr) #status experimental

Query Match      86.9%; Score 126; DB 2; Length 51;
Best Local Similarity 84.0%; Pred. No. 4.6e-12;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

DB 27 GGGVIGTISHECRMNSMOWFTCCS 51

RESULT 2
A58711
variacin precursor - Micrococcus varians
C:Species: Micrococcus varians
C:Date: 26-Jan-1998 #sequence_revision 30-Jan-1998 #text_change 24-Sep-1999
C:Accession: A58711
R:Pridmore, D.; Rekhif, N.; Pitet, A.C.; Surf, B.; Mollet, B.
Appl. Environ. Microbiol. 62, 1799-1802, 1996
A:Title: Variacin, a new lantibiotic-containing bacteriocin produced by Micrococcus v
A:Reference number: A58711; MUID:96209245
A:Accession: A58711

```

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-47 <PRT>
A:Cross-references: GB:X93303; NID:g1071698; PIDN:CA63706.1; PID:g1071699
A:Note: Parts of this sequence, including the amino end of the mature protein, were determined from the cDNA sequence.
C:Superfamily: unassigned lanthionine-containing peptides
C:Keywords: antibiotic; lanthionine
F:1-22/Domain: propeptide #status predicted <PRO>
F:23-47/Product: variacin #status predicted <MAT>
F:29-34/Cross-link: (25,35,6R)-3-methyl-lanthionine (Thr-Cys) #status predicted
F:31-45/Cross-link: sn-(25,6R)-lanthionine (Ser-Cys) #status predicted
F:38-46/Cross-link: sn-(25,6R)-lanthionine (Ser-Cys) #status predicted
F:44/Modified site: (2)-dehydrobutyrine (Thr) #status predicted

Query Match
Best Local Similarity 79.3%; Score 115; DB 2; Length 47;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db
1 GGGVIGTISHECRNMSWQFLTCSS 25
| ||| ||||| ||| ||| |||||
23 GSGVIPRTISHECRNMSWQFLTCSS 47

RESULT 3
JC6526
Lanthibiotic mutacin II precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 24-Sep-1999
C:Accession: JC6526
R:Woodruff, W.A.; Novak, J.; Caulfield, P.W.
Gene 206, 37-43, 1998
A:Title: Sequence analysis of muta and mutM genes involved in the biosynthesis of the lanthibiotic mutacin II precursor
A:Reference number: JC6526; MUID:98121310
A:Accession: JC6526
A:Molecule type: mRNA
A:Residues: 1-53 <MOO>
A:Cross-references: GB:40620; NID:g2853234; PIDN:AC38144.1; PID:g2853235
C:Genetics:
A:Gene: muta
C:Superfamily: unassigned lanthionine-containing peptides
C:Keywords: antibiotic; lanthionine
F:1-27/Domain: propeptide #status predicted <SIG>
F:28-53/Product: lanthibiotic mutacin II #status predicted <MAT>
F:36-41/Cross-link: (25,35,6R)-3-methyl-lanthionine (Thr-Cys) #status predicted
F:38-52/Cross-link: sn-(25,6R)-lanthionine (Ser-Cys) #status predicted
F:45-53/Cross-link: sn-(25,6R)-lanthionine (Ser-Cys) #status predicted
F:51/Modified site: (2)-dehydrobutyrine (Thr) #status predicted

Query Match
Best Local Similarity 72.4%; Score 105; DB 2; Length 53;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db
3 GVIOTISHECRNMSWQFLTC 24
||| ||| ||||| ||||| |||||
32 GVPVTSYECRNMSWQFLTC 53

RESULT 4
T09004
Streptococcin A-Ff22 precursor - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T09004; A58599
R:Hynes, W.L.; Perrella, J.J.; Tagg, J.R.
Appl. Environ. Microbiol. 59, 1969-1971, 1993
A:Title: Cloning of the gene encoding streptococcin A-Ff22, a novel lanthibiotic produced by Streptococcus pyogenes
A:Reference number: A58598; MUID:93319301
A:Accession: T09004
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-51 <HYN>

```

A:Cross-references: EMBL:AF026542; NID:g2502065; PIDN:AA892600.1; PID:g2502068
A:Experimental source: strain FF22
R:Jack, R.W.; Carne, A.; Metzger, J.; Stefanovic, S.; Sahl, H.G.; Jung, G.; Tagg, J.
Eur. J. Biochem. 220, 455-462, 1994
A:Title: Elucidation of the structure of SA-FF22, a lanthionine-containing antibacter
A:Reference number: A58599; MUID:94170794
A:Accession: A58599
A:Molecule type: protein
A:Residues: 26-51 <JMC>
A:Note: the lanthionine cross-links were not determined
C:Genetics:
A:Gene: scna
C:Keywords: antibiotic; lanthionine
F:1-25/Domain: propeptide #status predicted <SIG>
F:26-51/Product: streptococin A-FF22 #status experimental <MAT>
F:33-38/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status predicted
F:35-49/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status predicted
F:42-50/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status predicted
F:48/Modified site: (Z)-dehydrobutyrine (Thr) #status predicted

Query Match 69.0%; Score 100; DB 2; Length 51;
Best Local Similarity 69.6%; Pred. No. 3.9e-08;
Matches 16; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 GVIQTISHECRMNSWQFLFTCCS 25
||:||||| :||:|||||
Db 29 GVFKTISHCHLNTWAFATCCS 51

RESULT 5
T09005
variacin homolog scna1 precursor - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T09005
R:Hynes, W.L.; Ferretti, J.J.; Tagg, J.R.
Appl. Environ. Microbiol. 59, 1969-1971, 1993
A:Title: Cloning of the gene encoding streptococin A-FF22, a novel lantibiotic produ
A:Reference number: A58598; MUID:93319301
A:Accession: T09005
A:Status: translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-51 <HY>
A:Cross-references: EMBL:AF026542; NID:g2502065; PIDN:AA892601.1; PID:g2502069
C:Genetics:
A:Gene: scna1
C:Superfamily: unassigned lanthionine-containing peptides
C:Keywords: antibiotic; lanthionine
F:1-27/Domain: propeptide #status predicted <SIG>
F:28-51/Product: variacin homolog #status predicted <MAT>
F:34-39/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status predicted
F:36-50/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status predicted
F:43-51/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status predicted
F:49/Modified site: (Z)-dehydrobutyrine (Thr) #status predicted

Query Match 51.7%; Score 75; DB 2; Length 51;
Best Local Similarity 60.9%; Pred. No. 0.00023;
Matches 14; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 2 GGVITQISHECRMNSWQFLFTCC 24
|:||||| ||:|||||
Db 29 GGVNTISAECRMNSLQAFITCC 51

RESULT 6
T08839
polyprotein - marmoset hepatitis GB virus A
C:Species: marmoset hepatitis GB virus A
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000
C:Accession: T08839

R:Rifery, J.C.; Deadi, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.
J. Gen. Virol. 79, 41-45, 1998
A:Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.
A:Reference number: Z16486; MUID:98120818
A:Accession: T08839
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-2970 <ERK>
A:Cross-references: EMBL:AF023424; NID:g2828597; PIDN:AMC40501.1; PID:g2828598
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: polypeptide

Query Match	44.88;	Score 65;	DB 2;	Length 2970;
Best Local Similarity	48.08;	Pred. No. 0.33;		
Matches 12; Conservative	2;	Mismatches 11;	Indels 0;	Gaps 0;

QY 1 GGGVITQITISHECRMSWQFLTYCCS 25
 ||||: |||: || |||
 Db 14 GGGVLFPSKHKRCRVGDRELSNCCS 38

RESULT 7
T08841
polyprotein - douroucouli hepatitis GB virus A
C:Species: douroucouli hepatitis GB virus A
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #ext_change 17-Nov-2000
C:Accession: T08841
R:Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.
J. Gen. Virol. 79, 41-45, 1998
A:Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.
A:Reference number: Z16486; M01D:98120818
A:Accession: T08841
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-3005 <ERK>
A:Cross-references: EMBL:AF023425; NTD:g2828599; PIDN:MAC40502.1; PID:g2828600
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein

		37.9%	Score 55;	DB 2;	Length 3005;
Query Match		37.5%	Pred. No. 11;		
Best local Similarity					
Matches	9;	Conservative	3;	Mismatches	12;
				Indels	0;
				Gaps	0.
OY	2	GGVITQITSECCRMNSQWLFCCS	25		
		: :			
		: :			
Db	15	GAFLQPAASHACHAAGTYVLFNCSS	38		

```

RESULT      8
TILBP7
tail fiber protein - phage T7
C:Species: phage T7
C:Date: 13-Jun-1993 #sequence-Revision 13-Jun-1993 #text-change 23-Jul-1999
C:Accession: A04373; S4233
R:Dunn, J. J.; Thompson, K.
submitted to the Nucleic Acid Sequence Database, September 1982
A:Reference number: A94615
A:Accession: A04373
A:Molecule type: DNA
A:Residues: 1-553 <DUN>
R:Dunn, J. J.; Studler, F. W.
J. Mol. Biol. 166, 477-535, 1993
A:Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 ge
A:Reference number: S42283; MUID:83241725
A:Accession: S42333
A:Molecule type: DNA
A:Residues: 1-553 <DUN>
A:Cross-references: EMBL:V01146; NID:9431187; PIDN:CA24435.1; PID:915611
A:Note: the authors did not translate the codon for residue 1
:Genetics:
:Gene: 17

```

A;Map position: 86.70-90.85
C;Superfamily: phage T7 tail fiber protein
C;Keywords: tail fiber

Query Match	36.2%	Score 52.5;	DB 1;	length 553;
Best Local Similarity	38.7%;	Pred. No. 5.2;		
Matches 12;	Conservative 2;	Mismatches 8;	Indels 9;	Gaps 1;

```

QY      1  GGGVIQTISHE-----CRMNSWQLEF  22
          ||||| : : | | | | |
Db      479 GGGVSVTVSQDLRFRIWIKCANNSWNFFRT  509

```

RESULT 9
G98217
hypothetical protein AGR_L1365 [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: G98217
R:Goodner, B.; Hinkle, G.; Gatlung, S.; Miller, N.; Blanchard, M.; Quirollo, B.; Goldm
A.: Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2023-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: G98217
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-412 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89265.1; PID:gl5159094; GSPDB:GN00170
A:Genetics:
A:Gene: AGR_L1365
A:Map position: linear chromosome
C:Superfamily: Synecchocystis hypothetical protein slr0049

Query Match	35.28;	Score 51;	DB 2;	Length 412;
Best Local Similarity	53.38;	Pred. No. 6.7;		
Matches	8;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0;

Qy 2 GGVIQITISHECRMNS 16
||| |::|:| ||
Db 11 GGVAQVVAHKCAQNS 25

RESULT 10
AD3069
conserved hypothetical protein Atu4170 [imported] - Agrobacterium tumefaciens (strain
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence-revision 11-Jan-2002 #text-change 01-Feb-2002
C:Accession: AD3069
R:Wood, D.W.; Seubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo-
lerage, G.; Gillet, P.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McEl-
karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2333, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam-
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AD3069
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-412 <KUR>
A:Cross-references: GB:AE008689; PID:AL44970.1; PID:g17742626; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4170
A:Map position: linear chromosome
;Superfamily: Synchocystis hypothetical protein slr0049

Query Match	35.28; Score 51; DB 2; Length 412;
Best Local Similarity	53.38; Pred. No. 6.7;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGVYQRTSHRCRMS 16
||| |::| |
Db 11 GGVAQVVAHKCAQNS 25

RESULT 11
E96035
conserved hypothetical protein Smb21630 [imported] - *Sinorhizobium meliloti* (strain 1021)
C:Species: *Sinorhizobium meliloti*
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: E96035
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; M01D:21396508; PMID:11481431
A:Accession: E96035
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <KUR>
A:Cross-references: GB:AL591985; PIDN:GAC49949.1; PID:g15141437; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandebol, M.; Vorholter, F.J.; Weidner, S.; Welis, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; M01D:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Genome: Smb21630
A:Superfamily: *Synechocystis* hypothetical protein slr0049

Query Match 35.2%; Score 51; DB 2; Length 413;
Best Local Similarity 53.3%; Pred. No. 6.7;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGVYQRTSHRCRMS 16
||| |::| |
Db 11 GGVAQVVAHKCAQNS 25

RESULT 12
AF3450
cathoxynorspermidine dehydrogenase (EC 1.1.1.-) [imported] - *Brucella melitensis* (strain
C:Species: *Brucella melitensis*
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AF3450
R:DelVecchio, V.G.; Kaparakis, V.; Radkar, R.J.; Patra, G.; Muir, C.; Los, T.; Ivanova,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3450
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52769.1; PID:g17983603; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11588
A:Map position: I
C:Superfamily: *Synechocystis* hypothetical protein slr0049
C:Keywords: oxidoreductase

Query Match 35.2%; Score 51; DB 2; Length 413;
Best Local Similarity 53.3%; Pred. No. 6.7;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGVYQRTSHRCRMS 16
||| |::| |
Db 11 GGVAQVVAHKCAQNS 25

RESULT 13
S53457
dominant autoantigen gp 330 - rat (fragment)
C:Species: *Rattus norvegicus* (Norway rat)
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 18-Aug-2000
C:Accession: S53457
R:Okhadze, G.G.; Oleinikov, A.V.; Kanakas, J.J.; Makker, S.P.
Biochem. J. 305, 711-713, 1995
A:Title: Different molecular forms of rat kidney gp330, the dominant autoantigen of
A:Reference number: S53457; M01D:95151000
A:Accession: S53457
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1650 <JOK>
C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-bln
F:2-106/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:71-147/Domain: EGF homology <EG1>
F:153-188/Domain: EGF homology <EG2>
F:196-235/Domain: LDL receptor WYTD-containing repeat homology <YW01>
F:236-278/Domain: LDL receptor WYTD-containing repeat homology <YW02>
F:279-329/Domain: LDL receptor WYTD-containing repeat homology <YW03>
F:330-373/Domain: LDL receptor WYTD-containing repeat homology <YW04>
F:374-414/Domain: LDL receptor WYTD-containing repeat homology <YW05>
F:415-457/Domain: LDL receptor WYTD-containing repeat homology <YW06>
F:466-505/Domain: EGF homology <EG3>
F:509-545/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:550-586/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:595-631/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:636-672/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:679-715/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:720-755/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:760-794/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:799-833/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F:843-879/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F:884-921/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F:929-963/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F:971-1006/Domain: EGF homology <EG4>
F:1012-1048/Domain: EGF homology <EG5>
F:1055-1099/Domain: LDL receptor WYTD-containing repeat homology #status atypical <Y
F:1100-1154/Domain: LDL receptor WYTD-containing repeat homology #status atypical <Y
F:1155-1188/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F:1189-1232/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F:1234-1273/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F:1274-1316/Domain: LDL receptor WYTD-containing repeat homology <YW13>
F:1326-1359/Domain: EGF homology <EG6>

Query Match 34.5%; Score 50; DB 2; Length 1650;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 ISHCRMSWQ 18
:|||| |
Db 591 VSHCESNEMO 601

RESULT 14
TA7690
hypothetical protein T72E16.170 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: TA7690
R:Benes, V.; Murbach, E.; Drzonek, H.; Ansoorge, W.; Mewes, H.W.; Lemcke, K.; Mayer,
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24472

Query Match	34.1%	Score 49.5	DB 2	Length 606
Best Local Similarity	33.3%	Pred. No. 16		
Matches 11	Conservative 5	Mismatches 4	Indels 13	Gaps 1

```

QY      5  IQTISH-----ECRMSWQLETC 24
          |:::|          :::| | | |
Db      98  IETINHPQLVNETSDIMCKKVLRSFLATCC 130

```

RESULT	15
T39569	

Probable alpha-glucan synthase - fission yeast (schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: J39559
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Brown, D.; Churcher, C.M.
 submitted to the EMBL Data Library, March 1999
 A:Reference number: Z21864
 A:Accession: J39559
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-358 <WOO>
 A:Cross-references: EMBL:AL035637; PIDN:CAB38509.1; GSPDB:GN00067; SPDB:SPBC16D10.05
 A:Experimental source: strain 972h-; cosmid c16D10
 C:Genetics:
 A:Gene: SPDB:SPBC16D10.05
 A:Map position: 2
 A:Introns: 119/3

Query Match	33.8%	Score 49:	DB 2:	Length 2358:
Best Local Similarity	33.3%	Pred. No.	69:	
Matches	8:	Mismatches	11:	Indels 0:
		Conservative		Gaps 0:

QY	1	GGGVIQTISHECRMNSWQFLFTCC	24
		: : : :	
Db	2049	GAGVSWIVRACIVGFGQIWWACC	2072

Search completed: May 24, 2002, 09:52:40
Job time: 170 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:57:41 ; Search time 13.51 Seconds

(without alignments)
71.650 Million cell updates/sec

Title: US-09-913-763-3

Perfect score: 145

Sequence: 1 GGGVIGTISHCCRNMSQFLFTCCS 25

Scoring table: BLOSUM62

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	86.9	51	LANA_LACIA	P36499 lactococcus
2	100	69.0	51	LANA_STRPY	P36501 streptococc
3	100	69.0	51	LANB_STRPY	Q54957 streptococc
4	52.5	36.2	553	VTFP_BPT7	P03748 bacterioph
5	51	35.2	464	NORM_PASMU	Q9CMZ9 pasteurilla
6	50	34.5	353	VGLM_MCMVK	P52373 murine cyto
7	49	33.8	2358	MOKD_SCHPO	Q9Y719 schizosacch
8	47	32.4	347	DHAS_LEGPN	Q31219 legionella
9	46.5	32.1	110	CYOD_PSEPU	Q9WWR4 pseudomonas
10	46	31.7	359	INX2_SCHAM	Q9XYN1 schistocerc
11	45	31.0	352	GLN2_PRAAL	P20805 frankia aln
12	45	31.0	461	Y061_MYCGE	P47307 mycoplasma
13	45	31.0	634	Y015_MYCPN	P75094 mycoplasma
14	45	31.0	898	FAS2_SCHAM	P22648 schistocerc
15	44.5	30.7	186	CHS2_HUSTMA	P30599 ustilago ma
16	44	30.3	363	WNT5_HALRO	O15978 halocynthia
17	44	30.3	932	PKNI_CHLPN	Q92986 chlamydia p
18	44	30.3	997	ATST_HUMAN	Q9UKP4 homo sapien
19	43	29.7	48	LANA_STRSL	P36500 streptococc
20	43	29.7	308	GLS2_ECOLI	P77470 escherichia
21	43	29.7	688	YEO7_YEAST	P40050 saccharomyc
22	43	29.7	955	TSP4_XENLA	Q06441 xenopus lae
23	43	29.7	3898	POLG_HCVB	P21530 hog cholera
24	42	29.0	197	XYNA_SCHCO	P35809 schizophy11
25	42	29.0	331	VE05_VACCC	P21046 vaccinia vi
26	42	29.0	341	VE05_VACCD	Q01478 vaccinia vi
27	42	29.0	341	VE05_VACCV	P21606 vaccinia vi
28	42	29.0	341	VE05_VACV	Q01483 variola vir
29	42	29.0	564	Y61A_MYCPN	P75041 mycoplasma
30	42	29.0	1013	CHSA_MYEMI	P30584 emeticella
31	42	29.0	2482	VME_PIG	Q28833 sus scrofa
32	41.5	28.6	250	Y282_BUCAT	P57369 buchnera ap
33	41.5	28.6	514	T3RH_HAEIN	P44105 haemophilus

ALIGNMENTS

RESULT	1	LANA_LACIA	STANDARD;	PRT;	51 AA.
AC	P36499;				
DT	01-JUN-1994 (Rel. 29, Created)				
DT	01-JUN-1994 (Rel. 29, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Antibiotic lactacin 481 precursor (Lactococcus DR).				
GN	LCTA OR LCT OR LCNDRI.				
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).				
OC	Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				
OC	Lactococcus;				
OX	NCBI_Taxid:1360;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND STRUCTURE BY NMR.				
RC	STRATIN-CNRZ 481;				
RX	MEDLINE-93346379; PubMed-8344922;				
RA	Piard J.-C., Kuipers O.P., Rollet H.S., Desmazeud M.J., de Vos W.M.;				
RT	"Structure, organization, and expression of the lct gene for lactacin				
RT	481, a novel antibiotic produced by Lactococcus lactis.";				
RL	J. Biol. Chem. 268:16361-16368(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRATIN-ADRIA 851030;				
RX	MEDLINE-94288641; PubMed-8017945;				
RA	Rance A., Dufour A., le Pogam S., Thuaud D., Bourgeois C.M.,				
RT	Pennee J.P.;				
RT	"Cloning, expression, and nucleotide sequence of genes involved in				
RT	production of lactococcin DR, a bacteriocin from Lactococcus lactis				
RT	subsp. lactis.";				
RL	Appl. Environ. Microbiol. 60:1652-1657(1994).				
RN	[3]				
RP	SEQUENCE OF 25-31.				
RC	STRATIN-CNRZ 481;				
RA	Piard J.-C., Muriana P.M., Desmazeud M.J., Kienhammer T.R.;				
RT	"Purification and partial characterization of lactacin 481, a				
RT	lactinone-containing bacteriocin produced by Lactococcus lactis				
RT	subsp. lactis CNRZ 481.";				
RL	Appl. Environ. Microbiol. 58:279-284(1992).				
RN	[4]				
RP	THIOETHER BONDS.				
RX	MEDLINE-96350545; PubMed-8764998;				
RA	van den Hooven H.W., Lagerwert F.M., Heerma W., Haverkamp J.,				
RA	Piard J.-C., Halbers C.W., Slezan R.J., Kuipers O.P., Rollet H.S.;				
RT	"The structure of the antibiotic lactacin 481 produced by				
RT	Lactococcus lactis: location of the thioether bridges.";				
RL	FEBS Lett. 391:317-322(1996)				
CC	- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTHIBIOTIC)				
CC	ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF				
CC	LANTHIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL				
CC	CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS				
CC	TRANSMEMBRANE PORES. LACTICIN 481 IS A BROAD SPECTRUM BACTERIACIN				
CC	EXHIBITING ACTIVITY AGAINST A WIDE RANGE OF LACTIC ACID BACTERIA				
CC	AND C. TROPHOTRICUM.				
CC	- SUBUNIT: MONOMER OR HOMODIMER.				
CC	- PTM: MATURATION OF LANTHIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF				

34	41.5	28.6	697	1	LCEE_RAT
35	41.5	28.6	763	1	RGT2_YEAST
36	41.5	28.6	862	1	TP3B_HUMAN
37	41.5	28.6	1345	1	VCAP_HSVJ7
38	41.5	28.6	1485	1	CHS2_EXOTE
39	41	28.3	189	1	CHS2_XYDBA
40	41	28.3	231	1	VANR_ENTFC
41	41	28.3	291	1	DTCM_MOUSE
42	41	28.3	421	1	VOR3_FOWPV
43	41	28.3	556	1	RORA_HUMAN
44	41	28.3	571	1	PEN3_ADE02
45	41	28.3	571	1	PEN3_ADE05

P33124	rattus norv
Q12300	saccharomyc
O95985	homo sapien
P52347	human herpe
P30586	exophiala j
P30604	xylohypha b
Q06239	enterococcu
P01882	mus musculu
O72903	fowlpox vir
P35398	homo sapien
P03276	human adeno
P12538	human adeno

```
CC      THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE
CC      BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLLOCATION AND
CC      CLEAVAGE OF THE MODIFIED PRECURSOR.
CC      -I- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
CC      -----
CC      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X71410; CA50534.1; -.
DR      EMBL; U91581; AAC72257.1; -.
DR      PIR; S36019; S36019.
DR      PIR; A47342; A47342.
KW      Antibiotic; Bacteriocin; Lantibiotic.
FT      PROPEP          1       24
FT      CHAIN           25      51      LANTIBIOTIC LACTICIN 481.
FT      MOD_RES         33      33      D-ABU (AMINOBOUTYRIC ACID).
FT      MOD_RES         35      35      D-ALANINE.
FT      MOD_RES         42      42      D-ALANINE.
FT      MOD_RES         48      48      DHB (2,3-DIDEHYDROBUTYRINE).
FT      THIOETH         33      38      ABU-S-CYS (BETA-METHYLLANTHIONINE).
FT      THIOETH         35      49      ALA-S-CYS (LANTHIONINE).
FT      THIOETH         42      50      ALA-S-CIS (LANTHIONINE).
SQ      SEQUENCE        51 AA;  5677 MW;  44A3951AC2CEAE3D CRC64;

Query Match              86.9%; Score 126; DB 1; Length 51;
Best Local Similarity    84.0%; Pred. NO. 6.5e-12;
Matches   21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 GGGVIGITISHECRMNSMQFLFTCCS 25
Db      27 GSGVIHTISHECNMNSMQVFFTCCS 51

RESULT      2
LANA_STRPY ID LANA_STRPY STANDARD; PROT; 51 AA.
AC P36501;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lantibiotic streptococcin A-FP22 precursor (Antibacterial peptide
DE SA-FP22).
GN SCNA.
OS Streptococcus pyogenes.
OC Plasmid.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCB1_TaxId=1314;
RX [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FP22;
RX MEDLINE=93319301; PubMed=8328813;
RA Hynes W.L., Ferretti J.J., Tagg J.R.;
RT Cloning of the gene encoding Streptococcin A-FP22, a novel
RT lantibiotic produced by Streptococcus pyogenes, and determination of
RT its nucleotide sequence.
RL Appl. Environ. Microbiol. 59:1969-1971(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FP22;
RA McLaughlin R.E., Hynes W.L.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RC SEQUENCE OF 26-51.
RC STRAIN=FP22;
RX MEDLINE=94110794; PubMed=8125103;
RX Jack R.W., Carne A., Metzger J., Stefanovic S., Sahl H.-G., Jung G.,
```

Tagg J.R. "Elucidation of the structure of SA-PF22, a lanthionine-containing antibacterial peptide produced by *Streptococcus pyogenes* strain PF22.";

Bur. J. Biochem. 220:455-462(1994).

-1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTHIBIOTIC) ACTIVE ON CERTAIN GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF LANTHIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS TRANSMEMBRANE PORES.

-1- SUBCELLULAR LOCATION: EITHER CELL ASSOCIATED OR IN A RELEASED EXTRACELLULAR FORM.

-1- PM: MATURATION OF LANTHIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND CLEAVAGE OF THE MODIFIED PRECURSOR.

-1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTHIBIOTICS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way used to modify and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF026542; AAB92600.1; -

DR Antibiotic; Bacteriocin; Lanthibiotic; Plasmid.

KW PROPEP 1 25

FT CHAIN 26 51 LANTHIBIOTIC STREPTOCOCCIN A-PF22.

FT MOD_RES 33 33 D-ABU (AMINOBTYRIC ACID).

FT MOD_RES 35 35 D-ALANINE.

FT MOD_RES 42 42 D-ABU (AMINOBTYRIC ACID).

FT MOD_RES 48 48 DHB (2,3-DIDEHYDROBTYRINE).

FT THIOETH 235 238 ALA-S-CYS (LANTHIONINE).

FT THIOETH 242 249 ABU-S-CYS (BETA-METHYLLANTHIONINE).

FT THIOETH 248 250 ABU-S-CYS (BETA-METHYLLANTHIONINE).

SO SEQUENCE 51 AA; 5666 MW; 77E378C7A1B9DAAC CRC64;

Query Match 69.0%; Score 100; DB 1; Length 51;
Best Local Similarity 69.6%; Pred. No. 4e-08;
Matches 16; Conservative 3; Mismatches 4; Indels 0; Gaps 0.

OY 3 GVIGTISHCRRMSWQFLTCCS 25
||:|||||:|:|:| |||||

DB 29 GVGFTHSHECHLNTWAFATCCS 51

RESULT 3.

LANB_STRPY STANDARD; PRT; 51 AA.

AC 054957;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Lanthibiotic streptococcin A-M49 precursor (Antibacterial peptide A-M49).

GN SCNA' AND SCNA'.

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

CC Streptococcus.

OX NCBI_Taxid-1314;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GT9538;

RX MEDLINE=95085281; PubMed=7993103;

RT "Hyne W.L., Friend V.L., Ferretti J.J.;"

RT "Duplication of the lantibiotic structural gene in M-type 49 group A streptococcus strains producing streptococcin A-M49.";

RL Appl. Environ. Microbiol. 60:4207-4209(1994).

-1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTHIBIOTIC) ACTIVE ON CERTAIN GRAM-POSITIVE BACTERIA. THE BACTERICIDAL

CC ACTIVITY OF LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED
 CC BACTERIAL CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF
 CC AQUEOUS TRANSMEMBRANE PORES (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: EITHER CELL ASSOCIATED OR IN A RELEASED
 CC EXTRACELLULAR FORM (BY SIMILARITY).
 CC -1- PTM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF
 CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE
 CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLLOCATION AND
 CC CLEAVAGE OF THE MODIFIED PRECURSOR (BY SIMILARITY).
 CC -1- MISCELLANEOUS: THERE ARE TWO GENES CODING FOR LANTIBIOTIC
 CC STREPTOCOCцин A-M49, THEIR CODING SEQUENCE ONLY DIFFERS IN THE
 CC PROPEPTIDE REGION.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L36235; AAA62362.1; -
 DR EMBL: L36235; AAA62361.1; -
 KW Antibiotic; Bacteriocin; Lantibiotic; Multigene family.
 FT PROPEP 1 25 LANTIBIOTIC STREPTOCOCцин A-M49.
 FT PEPTIDE 26 51 BY SIMILARITY.
 FT MOD_RES 33 33 D-ABU (AMINOISOBUTYRIC ACID)
 FT MOD_RES 35 35 (BY SIMILARITY).
 FT MOD_RES 42 42 D-ALANINE (BY SIMILARITY).
 FT MOD_RES 48 48 D-ABU (AMINOISOBUTYRIC ACID)
 FT MOD_RES 49 49 (BY SIMILARITY).
 FT MOD_RES 50 50 DHB (2,3-DIDEHYDROBUTYRINE)
 FT MOD_RES 51 51 (BY SIMILARITY).
 FT THIOETH 35 38 ALA-S-CYS (LANTHIONINE) (BY SIMILARITY).
 FT THIOETH 42 49 ABU-S-CYS (BETA-METHYLLANTHIONINE)
 FT THIOETH 48 50 ABU-S-CYS (BETA-METHYLLANTHIONINE)
 FT THIOETH 50 50 (BY SIMILARITY).
 FT THIOETH 51 51 T -> E (IN SCNA'').
 FT THIOETH 52 52 EH -> NN (IN SCNA'').
 FT THIOETH 53 53 I -> V (IN SCNA'').
 FT THIOETH 54 54 VARIANT 4 5
 FT THIOETH 55 55 VARIANT 7 7
 FT THIOETH 56 56 VARIANT 7 7
 FT THIOETH 57 57 SEQUENCE 51 AA; 5690 MW; F12124BA13E17FD1 CRC64;
 FT THIOETH 58 58
 FT THIOETH 59 59
 FT THIOETH 60 60
 FT THIOETH 61 61
 FT THIOETH 62 62
 FT THIOETH 63 63
 FT THIOETH 64 64
 FT THIOETH 65 65
 FT THIOETH 66 66
 FT THIOETH 67 67
 FT THIOETH 68 68
 FT THIOETH 69 69
 FT THIOETH 70 70
 FT THIOETH 71 71
 FT THIOETH 72 72
 FT THIOETH 73 73
 FT THIOETH 74 74
 FT THIOETH 75 75
 FT THIOETH 76 76
 FT THIOETH 77 77
 FT THIOETH 78 78
 FT THIOETH 79 79
 FT THIOETH 80 80
 FT THIOETH 81 81
 FT THIOETH 82 82
 FT THIOETH 83 83
 FT THIOETH 84 84
 FT THIOETH 85 85
 FT THIOETH 86 86
 FT THIOETH 87 87
 FT THIOETH 88 88
 FT THIOETH 89 89
 FT THIOETH 90 90
 FT THIOETH 91 91
 FT THIOETH 92 92
 FT THIOETH 93 93
 FT THIOETH 94 94
 FT THIOETH 95 95
 FT THIOETH 96 96
 FT THIOETH 97 97
 FT THIOETH 98 98
 FT THIOETH 99 99
 FT THIOETH 100 100
 FT THIOETH 101 101
 FT THIOETH 102 102
 FT THIOETH 103 103
 FT THIOETH 104 104
 FT THIOETH 105 105
 FT THIOETH 106 106
 FT THIOETH 107 107
 FT THIOETH 108 108
 FT THIOETH 109 109
 FT THIOETH 110 110
 FT THIOETH 111 111
 FT THIOETH 112 112
 FT THIOETH 113 113
 FT THIOETH 114 114
 FT THIOETH 115 115
 FT THIOETH 116 116
 FT THIOETH 117 117
 FT THIOETH 118 118
 FT THIOETH 119 119
 FT THIOETH 120 120
 FT THIOETH 121 121
 FT THIOETH 122 122
 FT THIOETH 123 123
 FT THIOETH 124 124
 FT THIOETH 125 125
 FT THIOETH 126 126
 FT THIOETH 127 127
 FT THIOETH 128 128
 FT THIOETH 129 129
 FT THIOETH 130 130
 FT THIOETH 131 131
 FT THIOETH 132 132
 FT THIOETH 133 133
 FT THIOETH 134 134
 FT THIOETH 135 135
 FT THIOETH 136 136
 FT THIOETH 137 137
 FT THIOETH 138 138
 FT THIOETH 139 139
 FT THIOETH 140 140
 FT THIOETH 141 141
 FT THIOETH 142 142
 FT THIOETH 143 143
 FT THIOETH 144 144
 FT THIOETH 145 145
 FT THIOETH 146 146
 FT THIOETH 147 147
 FT THIOETH 148 148
 FT THIOETH 149 149
 FT THIOETH 150 150
 FT THIOETH 151 151
 FT THIOETH 152 152
 FT THIOETH 153 153
 FT THIOETH 154 154
 FT THIOETH 155 155
 FT THIOETH 156 156
 FT THIOETH 157 157
 FT THIOETH 158 158
 FT THIOETH 159 159
 FT THIOETH 160 160
 FT THIOETH 161 161
 FT THIOETH 162 162
 FT THIOETH 163 163
 FT THIOETH 164 164
 FT THIOETH 165 165
 FT THIOETH 166 166
 FT THIOETH 167 167
 FT THIOETH 168 168
 FT THIOETH 169 169
 FT THIOETH 170 170
 FT THIOETH 171 171
 FT THIOETH 172 172
 FT THIOETH 173 173
 FT THIOETH 174 174
 FT THIOETH 175 175
 FT THIOETH 176 176
 FT THIOETH 177 177
 FT THIOETH 178 178
 FT THIOETH 179 179
 FT THIOETH 180 180
 FT THIOETH 181 181
 FT THIOETH 182 182
 FT THIOETH 183 183
 FT THIOETH 184 184
 FT THIOETH 185 185
 FT THIOETH 186 186
 FT THIOETH 187 187
 FT THIOETH 188 188
 FT THIOETH 189 189
 FT THIOETH 190 190
 FT THIOETH 191 191
 FT THIOETH 192 192
 FT THIOETH 193 193
 FT THIOETH 194 194
 FT THIOETH 195 195
 FT THIOETH 196 196
 FT THIOETH 197 197
 FT THIOETH 198 198
 FT THIOETH 199 199
 FT THIOETH 200 200
 FT THIOETH 201 201
 FT THIOETH 202 202
 FT THIOETH 203 203
 FT THIOETH 204 204
 FT THIOETH 205 205
 FT THIOETH 206 206
 FT THIOETH 207 207
 FT THIOETH 208 208
 FT THIOETH 209 209
 FT THIOETH 210 210
 FT THIOETH 211 211
 FT THIOETH 212 212
 FT THIOETH 213 213
 FT THIOETH 214 214
 FT THIOETH 215 215
 FT THIOETH 216 216
 FT THIOETH 217 217
 FT THIOETH 218 218
 FT THIOETH 219 219
 FT THIOETH 220 220
 FT THIOETH 221 221
 FT THIOETH 222 222
 FT THIOETH 223 223
 FT THIOETH 224 224
 FT THIOETH 225 225
 FT THIOETH 226 226
 FT THIOETH 227 227
 FT THIOETH 228 228
 FT THIOETH 229 229
 FT THIOETH 230 230
 FT THIOETH 231 231
 FT THIOETH 232 232
 FT THIOETH 233 233
 FT THIOETH 234 234
 FT THIOETH 235 235
 FT THIOETH 236 236
 FT THIOETH 237 237
 FT THIOETH 238 238
 FT THIOETH 239 239
 FT THIOETH 240 240
 FT THIOETH 241 241
 FT THIOETH 242 242
 FT THIOETH 243 243
 FT THIOETH 244 244
 FT THIOETH 245 245
 FT THIOETH 246 246
 FT THIOETH 247 247
 FT THIOETH 248 248
 FT THIOETH 249 249
 FT THIOETH 250 250
 FT THIOETH 251 251
 FT THIOETH 252 252
 FT THIOETH 253 253
 FT THIOETH 254 254
 FT THIOETH 255 255
 FT THIOETH 256 256
 FT THIOETH 257 257
 FT THIOETH 258 258
 FT THIOETH 259 259
 FT THIOETH 260 260
 FT THIOETH 261 261
 FT THIOETH 262 262
 FT THIOETH 263 263
 FT THIOETH 264 264
 FT THIOETH 265 265
 FT THIOETH 266 266
 FT THIOETH 267 267
 FT THIOETH 268 268
 FT THIOETH 269 269
 FT THIOETH 270 270
 FT THIOETH 271 271
 FT THIOETH 272 272
 FT THIOETH 273 273
 FT THIOETH 274 274
 FT THIOETH 275 275
 FT THIOETH 276 276
 FT THIOETH 277 277
 FT THIOETH 278 278
 FT THIOETH 279 279
 FT THIOETH 280 280
 FT THIOETH 281 281
 FT THIOETH 282 282
 FT THIOETH 283 283
 FT THIOETH 284 284
 FT THIOETH 285 285
 FT THIOETH 286 286
 FT THIOETH 287 287
 FT THIOETH 288 288
 FT THIOETH 289 289
 FT THIOETH 290 290
 FT THIOETH 291 291
 FT THIOETH 292 292
 FT THIOETH 293 293
 FT THIOETH 294 294
 FT THIOETH 295 295
 FT THIOETH 296 296
 FT THIOETH 297 297
 FT THIOETH 298 298
 FT THIOETH 299 299
 FT THIOETH 300 300
 FT THIOETH 301 301
 FT THIOETH 302 302
 FT THIOETH 303 303
 FT THIOETH 304 304
 FT THIOETH 305 305
 FT THIOETH 306 306
 FT THIOETH 307 307
 FT THIOETH 308 308
 FT THIOETH 309 309
 FT THIOETH 310 310
 FT THIOETH 311 311
 FT THIOETH 312 312
 FT THIOETH 313 313
 FT THIOETH 314 314
 FT THIOETH 315 315
 FT THIOETH 316 316
 FT THIOETH 317 317
 FT THIOETH 318 318
 FT THIOETH 319 319
 FT THIOETH 320 320
 FT THIOETH 321 321
 FT THIOETH 322 322
 FT THIOETH 323 323
 FT THIOETH 324 324
 FT THIOETH 325 325
 FT THIOETH 326 326
 FT THIOETH 327 327
 FT THIOETH 328 328
 FT THIOETH 329 329
 FT THIOETH 330 330
 FT THIOETH 331 331
 FT THIOETH 332 332
 FT THIOETH 333 333
 FT THIOETH 334 334
 FT THIOETH 335 335
 FT THIOETH 336 336
 FT THIOETH 337 337
 FT THIOETH 338 338
 FT THIOETH 339 339
 FT THIOETH 340 340
 FT THIOETH 341 341
 FT THIOETH 342 342
 FT THIOETH 343 343
 FT THIOETH 344 344
 FT THIOETH 345 345
 FT THIOETH 346 346
 FT THIOETH 347 347
 FT THIOETH 348 348
 FT THIOETH 349 349
 FT THIOETH 350 350
 FT THIOETH 351 351
 FT THIOETH 352 352
 FT THIOETH 353 353
 FT THIOETH 354 354
 FT THIOETH 355 355
 FT THIOETH 356 356
 FT THIOETH 357 357
 FT THIOETH 358 358
 FT THIOETH 359 359
 FT THIOETH 360 360
 FT THIOETH 361 361
 FT THIOETH 362 362
 FT THIOETH 363 363
 FT THIOETH 364 364
 FT THIOETH 365 365
 FT THIOETH 366 366
 FT THIOETH 367 367
 FT THIOETH 368 368
 FT THIOETH 369 369
 FT THIOETH 370 370
 FT THIOETH 371 371
 FT THIOETH 372 372
 FT THIOETH 373 373
 FT THIOETH 374 374
 FT THIOETH 375 375
 FT THIOETH 376 376
 FT THIOETH 377 377
 FT THIOETH 378 378
 FT THIOETH 379 379
 FT THIOETH 380 380
 FT THIOETH 381 381
 FT THIOETH 382 382
 FT THIOETH 383 383
 FT THIOETH 384 384
 FT THIOETH 385 385
 FT THIOETH 386 386
 FT THIOETH 387 387
 FT THIOETH 388 388
 FT THIOETH 389 389
 FT THIOETH 390 390
 FT THIOETH 391 391
 FT THIOETH 392 392
 FT THIOETH 393 393
 FT THIOETH 394 394
 FT THIOETH 395 395
 FT THIOETH 396 396
 FT THIOETH 397 397
 FT THIOETH 398 398
 FT THIOETH 399 399
 FT THIOETH 400 400
 FT THIOETH 401 401
 FT THIOETH 402 402
 FT THIOETH 403 403
 FT THIOETH 404 404
 FT THIOETH 405 405
 FT THIOETH 406 406
 FT THIOETH 407 407
 FT THIOETH 408 408
 FT THIOETH 409 409
 FT THIOETH 410 410
 FT THIOETH 411 411
 FT THIOETH 412 412
 FT THIOETH 413 413
 FT THIOETH 414 414
 FT THIOETH 415 415
 FT THIOETH 416 416
 FT THIOETH 417 417
 FT THIOETH 418 418
 FT THIOETH 419 419
 FT THIOETH 420 420
 FT THIOETH 421 421
 FT THIOETH 422 422
 FT THIOETH 423 423
 FT THIOETH 424 424
 FT THIOETH 425 425
 FT THIOETH 426 426
 FT THIOETH 427 427
 FT THIOETH 428 428
 FT THIOETH 429 429
 FT THIOETH 430 430
 FT THIOETH 431 431
 FT THIOETH 432 432
 FT THIOETH 433 433
 FT THIOETH 434 434
 FT THIOETH 435 435
 FT THIOETH 436 436
 FT THIOETH 437 437
 FT THIOETH 438 438
 FT THIOETH 439 439
 FT THIOETH 440 440
 FT THIOETH 441 441
 FT THIOETH 442 442
 FT THIOETH 443 443
 FT THIOETH 444 444
 FT THIOETH 445 445
 FT THIOETH 446 446
 FT THIOETH 447 447
 FT THIOETH 448 448
 FT THIOETH 449 449
 FT THIOETH 450 450
 FT THIOETH 451 451
 FT THIOETH 452 452
 FT THIOETH 453 453
 FT THIOETH 454 454
 FT THIOETH 455 455
 FT THIOETH 456 456
 FT THIOETH 457 457
 FT THIOETH 458 458
 FT THIOETH 459 459
 FT THIOETH 460 460
 FT THIOETH 461 461
 FT THIOETH 462 462
 FT THIOETH 463 463
 FT THIOETH 464 464
 FT THIOETH 465 465
 FT THIOETH 466 466
 FT THIOETH 467 467
 FT THIOETH 468 468
 FT THIOETH 469 469
 FT THIOETH 470 470
 FT THIOETH 471 471
 FT THIOETH 472 472
 FT THIOETH 473 473
 FT THIOETH 474 474
 FT THIOETH 475 475
 FT THIOETH 476 476
 FT THIOETH 477 477
 FT THIOETH 478 478
 FT THIOETH 479 479
 FT THIOETH 480 480
 FT THIOETH 481 481
 FT THIOETH 482 482
 FT THIOETH 483 483
 FT THIOETH 484 484
 FT THIOETH 485 485
 FT THIOETH 486 486
 FT THIOETH 487 487
 FT THIOETH 488 488
 FT THIOETH 489 489
 FT THIOETH 490 490
 FT THIOETH 491 491
 FT THIOETH 492 492
 FT THIOETH 493 493
 FT THIOETH 494 494
 FT THIOETH 495 495
 FT THIOETH 496 496
 FT THIOETH 497 497
 FT THIOETH 498 498
 FT THIOETH 499 499
 FT THIOETH 500 500
 FT THIOETH 501 501
 FT THIOETH 502 502
 FT THIOETH 503 503
 FT THIOETH 504 504
 FT THIOETH 505 505
 FT THIOETH 506 506
 FT THIOETH 507 507
 FT THIOETH 508 508
 FT THIOETH 509 509
 FT THIOETH 510 510
 FT THIOETH 511 511
 FT THIOETH 512 512
 FT THIOETH 513 513
 FT THIOETH 514 514
 FT THIOETH 515 515
 FT THIOETH 516 516
 FT THIOETH 517 517
 FT THIOETH 518 518
 FT THIOETH 519 519
 FT THIOETH 520 520
 FT THIOETH 521 521
 FT THIOETH 522 522
 FT THIOETH 523 523
 FT THIOETH 524 524
 FT THIOETH 525 525
 FT THIOETH 526 526
 FT THIOETH 527 527
 FT THIOETH 528 528
 FT THIOETH 529 529
 FT THIOETH 530 530
 FT THIOETH 531 531
 FT THIOETH 532 532
 FT THIOETH 533 533
 FT THIOETH 534 534
 FT THIOETH 535 535
 FT THIOETH 536 536
 FT THIOETH 537 537
 FT THIOETH 538 538
 FT THIOETH 539 539
 FT THIOETH 540 540
 FT THIOETH 541 541
 FT THIOETH 542 542
 FT THIOETH 543 543
 FT THIOETH 544 544
 FT THIOETH 545 545
 FT THIOETH 546 546
 FT THIOETH 547 547
 FT THIOETH 548 548
 FT THIOETH 549 549
 FT THIOETH 550 550
 FT THIOETH 551 551
 FT THIOETH 552 552
 FT THIOETH 553 553
 FT THIOETH 554 554
 FT THIOETH 555 555
 FT THIOETH 556 556
 FT THIOETH 557 557
 FT THIOETH 558 558
 FT THIOETH 559 559
 FT THIOETH 560 560
 FT THIOETH 561 561
 FT THIOETH 562 562
 FT THIOETH 563 563
 FT THIOETH 564 564
 FT THIOETH 565 565
 FT THIOETH 566 566
 FT THIOETH 567 567
 FT THIOETH 568 568
 FT THIOETH 569 569
 FT THIOETH 570 570
 FT THIOETH 571 571
 FT THIOETH 572 572
 FT THIOETH 573 573
 FT THIOETH 574 574
 FT THIOETH 575 575
 FT THIOETH 576 576
 FT THIOETH 577 577
 FT THIOETH 578 578
 FT THIOETH 579 579
 FT THIOETH 580 580
 FT THIOETH 581 581
 FT THIOETH 582 582
 FT THIOETH 583 583
 FT THIOETH 584 584
 FT THIOETH 585 585
 FT THIOETH 586 586
 FT THIOETH 587 587
 FT THIOETH 588 588
 FT THIOETH 589 589
 FT THIOETH 590 590
 FT THIOETH 591 591
 FT THIOETH 592 592
 FT THIOETH 593 593
 FT THIOETH 594 594
 FT THIOETH 595 595
 FT THIOETH 596 596
 FT THIOETH 597 597
 FT THIOETH 598 598
 FT THIOETH 599 599
 FT THIOETH 600 600
 FT THIOETH 601 601
 FT THIOETH 602 602
 FT THIOETH 603 603
 FT THIOETH 604 604
 FT THIOETH 605 605
 FT THIOETH 606 606
 FT THIOETH 607 607
 FT THIOETH 608 608
 FT THIOETH 609 609
 FT THIOETH 610 610
 FT THIOETH 611 611
 FT THIOETH 612 612
 FT THIOETH 613 613
 FT THIOETH 614 614
 FT THIOETH 615 615
 FT THIOETH 616 616
 FT THIOETH 617 617
 FT THIOETH 618 618
 FT THIOETH 619 619
 FT THIOETH 620 620
 FT THIOETH 621 621
 FT THIOETH 622 622
 FT THIOETH 623 623
 FT THIOETH 624 624
 FT THIOETH 625 625
 FT THIOETH 626 626
 FT THIOETH 627 627
 FT THIOETH 628 628
 FT THIOETH 629 629
 FT THIOETH 630 630
 FT THIOETH 631 631
 FT THIOETH 632 632
 FT THIOETH 633 633
 FT THIOETH 634 634
 FT THIOETH 635 635
 FT THIOETH 636 636
 FT THIOETH 637 637
 FT THIOETH 638 638
 FT THIOETH 639 639
 FT THIOETH 640 640
 FT THIOETH 641 641
 FT THIOETH 642 642
 FT THIOETH 643 643
 FT THIOETH 644 644
 FT THIOETH 645 645
 FT THIOETH 646 646
 FT THIOETH 647 647
 FT THIOETH 648 648
 FT THIOETH 649 649
 FT THIOETH 650 650
 FT THIOETH 651 651
 FT THIOETH 652 652
 FT THIOETH 653 653
 FT THIOETH 654 654
 FT THIOETH 655 655
 FT THIOETH 656 656
 FT THIOETH 657 657
 FT THIOETH 658 658
 FT THIOETH 659 659
 FT THIOETH 660 660
 FT THIOETH 661 661
 FT THIOETH 662 662
 FT THIOETH 663 663
 FT THIOETH 664 664
 FT THIOETH 665 665
 FT THIOETH 666 666
 FT THIOETH 667 667
 FT THIOETH 668 668
 FT THIOETH 669 669
 FT THIOETH 670 670
 FT THIOETH 671 671
 FT THIOETH 672 672
 FT THIOETH 673 673
 FT THIOETH 674 674
 FT THIOETH 675 675
 FT THIOETH 676 676
 FT THIOETH 677 677
 FT THIOETH 678 678
 FT THIOETH 679 679
 FT THIOETH 680 680
 FT THIOETH 681 681
 FT THIOETH 682 682
 FT THIOETH 683 683
 FT THIOETH 684 684
 FT THIOETH 685 685
 FT THIOETH 686 686
 FT THIOETH 687 687
 FT THIOETH 688 688
 FT THIOETH 689 689
 FT THIOETH 690 690
 FT THIOETH 691 691
 FT THIOETH 692 692
 FT THIOETH 693 693
 FT THIOETH 694 694
 FT THIOETH 695 695
 FT THIOETH 696 696
 FT THIOETH 697 697
 FT THIOETH 698 698
 FT THIOETH 699 699
 FT THIOETH 700 700
 FT THIOETH 701 701
 FT THIOETH 702 702
 FT THIOETH 703 703
 FT THIOETH 704 704
 FT THIOETH 705 705
 FT THIOETH 706 706
 FT THIOETH 707 707
 FT THIOETH 708 708
 FT THIOETH 709 709
 FT THIOETH 710 710
 FT THIOETH 711 711
 FT THIOETH 712 712
 FT THIOETH 713 713
 FT THIOETH 714 714
 FT THIOETH 715 715
 FT THIOETH 716 716
 FT THIOETH 717 717
 FT THIOETH 718 718
 FT THIOETH 719 719
 FT THIOETH 720 720
 FT THIOETH 721 721
 FT THIOETH 722 722
 FT THIOETH 723 723
 FT THIOETH 724 724
 FT THIOETH 725 725
 FT THIOETH 726 726
 FT THIOETH 727 727
 FT THIOETH 728 728
 FT THIOETH 729 729
 FT THIOETH 730 730
 FT THIOETH 731 731
 FT THIOETH 732 732
 FT THIOETH 733 733
 FT THIOETH 734 734
 FT THIOETH 735 735
 FT THIOETH 736 736
 FT THIOETH 737 737
 FT THIOETH 738 738
 FT THIOETH 739 739
 FT THIOETH 740 740
 FT THIOETH 741 741
 FT THIOETH 742 742
 FT THIOETH 743 743
 FT THIOETH 744 744
 FT THIOETH 745 745
 FT THIOETH 746 746
 FT THIOETH 747 747
 FT THIOETH 748 748
 FT THIOETH 749 749
 FT THIOETH 750 750
 FT THIOETH 751 751
 FT THIOETH 752 752
 FT THIOETH 753 753
 FT THIOETH 754 754
 FT THIOETH 755 755
 FT THIOETH 756 756
 FT THIOETH 757 757

FT TRANSMEM 285 305 POTENTIAL.
 FT TRANSMEM 321 341 POTENTIAL.
 FT TRANSMEM 355 375 POTENTIAL.
 FT TRANSMEM 391 411 POTENTIAL.
 FT TRANSMEM 423 443 POTENTIAL.
 SQ SEQUENCE 464 AA: 50666 MW: A9945110F05A61C6 CRC64:

Query Match 35.2%; Score 51; DB 1; Length 464;
 Best Local Similarity 39.1%; Pred. No. 3.7;
 Matches 9; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 1 GGGVYQITISHECRNMSQWFLFTC 23
 Db 82 GSGKRODIAHQVROGIMILFSC 104

RESULT 6
 VGLM_MCMVK STANDARD; PRT; 353 AA.

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Glycoprotein M.
 GN GM OR U1100.
 OS Murine cytomegalovirus (strain K181).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Muromegalovirus.
 OX NCBI_Taxid=69156;

RA SEQUENCE FROM N.A.
 RX MEDLINE=96068827; PubMed=7595401;
 RA Scalzo A.A., Forbes C.A., Davis-Poynter N.J., Farrell H.E.,
 Lyons P.A.;
 RT "DNA sequence and transcriptional analysis of the glycoprotein M gene
 of murine cytomegalovirus."

RL J. Gen. Virol. 76:2895-2901(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- MISCELLANEOUS: ALSO SEQUENCED IN STRAINS G4, G6, K17B, K17E, K29
 CC AND N1. THE SEQUENCE IS IDENTICAL.

CC -1- SIMILARITY: NO OTHER HERPESVIRUSES GLYCOPROTEIN M

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L41088; AAC13736.1; -
 DR EMBL: L41089; AAC05166.1; -
 DR EMBL: L41090; AAC05167.1; -
 DR EMBL: L41091; AAC05168.1; -
 DR EMBL: L41092; AAC05169.1; -
 DR EMBL: L41093; AAC05170.1; -
 DR EMBL: L41094; AAC05171.1; -
 DR InterPro: IPR00785; Herpes_glycop.
 DR Pfam: PF01528; Herpes_glycop.1.
 DR PRINTS: PR00333; HSVINTGRAMP.

KW Transmembrane; Glycoprotein.
 FT TRANSMEM 28 48 POTENTIAL.
 FT TRANSMEM 83 103 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 211 231 POTENTIAL.
 FT TRANSMEM 241 261 POTENTIAL.
 FT TRANSMEM 271 291 POTENTIAL.
 FT TRANSMEM 305 325 POTENTIAL.
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 353 AA: 40106 MW: 021F1380F3E6F428 CRC64:

Query Match 34.5%; Score 50; DB 1; Length 353;
 Best Local Similarity 40.9%; Pred. No. 4.1;
 Matches 9; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

OY 5 IQTISHECRMN--SQWFLFTCC 24
 Db 6 VMTLSHVDRMNLRTWTMAIACC 27

RESULT 7
 MOKD_SCHPO STANDARD; PRT; 2358 AA.

AC 09Y719; 094638;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cell wall alpha-1,3-glucan synthase mok13 (EC 2.4.1.183).
 GN MOK13 OR SPBC16D10.05.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_Taxid=4896;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RA Katayama S., Dai H., Arellano M., Perez P., Toda T.;
 RT "Fission yeast alpha-glucan synthase Mok1 localizes closely with actin
 RT and play a role essential for cell morphogenesis and protein kinase C
 RT function."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=972;
 RA Wood V., Rajadream M.A., Barrell B.G., Brown D., Churcher C.M.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: UDP-glucose + [alpha-D-glucosyl-(1,3)](N) -
 CC UDP + [alpha-D-glucosyl-(1,3)](N+1).

CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB018382; BAA76559.1; -
 DR EMBL: AL035637; CAB38509.1; -
 DR InterPro: IPR000461; Alpha_amylase.
 DR Pfam: PF00128; alpha-amylase.1.
 KW Cell wall; Transferase; Glycosyltransferase.
 FT CONFLICT 120 120 V -> VRRVWLCSLTNKKV (IN REF. 1).
 SQ SEQUENCE 2358 AA: 269192 MW: A175577C9D8AD731 CRC64:

Query Match 33.8%; Score 49; DB 1; Length 2358;
 Best Local Similarity 33.3%; Pred. No. 30;
 Matches 8; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

OY 1 GGGVYQITISHECRNMSQWFLFTC 24
 Db 2049 GAGVAVSWIVRACIVGFGQIWAAC 2072

RESULT 8
 DHAS_LEGPN STANDARD; PRT; 347 AA.
 ID DHAS_LEGPN
 AC 031219;

DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11) (ASA dehydrogenase)
 DE (ASADH).
 GN ASD.
 OS Legionella pneumophila.
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
 OC Legionellaceae; Legionella.
 OX NCBI_TaxID=446;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AA100;
 RA MEDLINE=96234011; PubMed=9573067;
 RX Harb O.S., Abu Kwaik Y.;
 RT "Identification of the aspartate-beta-semialdehyde dehydrogenase gene
 of Legionella pneumophila and characterization of a null mutant."
 RT Infect. Immun. 66:1898-1903(1998).
 RL
 CC -1- CATALYTIC ACTIVITY: L-aspartate-4-semialdehyde + phosphate +
 CC NADP(+) = L-4-aspartyl phosphate + NADPH.
 CC -1- PATHWAY: SECOND STEP IN THE COMMON BIOSYNTHETIC PATHWAY LEADING
 CC FROM ASP TO THE CELL WALL. PRECURSOR MESO-DIAMINOPIMELATE, TO LYS,
 CC TO MET. TO ILE AND TO THR.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ASPARTATE-SEMIALDEHYDE DEHYDROGENASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF034213; AAC46292.1;
 DR HSSP: P00353; 1BRM.
 DR InterPro: IPR000519; Asp_semiald_dh.
 DR Pfam: PF01118; Semialdhyde_dh.
 DR Pfam: PF02774; Semialdhyde_dhc.1.
 DR PROSITE: PS01103; ASD; FALSE_NEG.
 DR Oxidoreductase; NADP; Diaminopimelate biosynthesis;
 KM Lysine biosynthesis.
 KM ACT_SITE 132 132 BY SIMILARITY.
 FT ACT_SITE 132 132
 SQ SEQUENCE 347 AA; 37714 MM; 33D992101D707943 CRC64;

 Query Match 32.4%; Score 47; DB 1; Length 347;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 OY 2 GGVITISHECRMSM 17
 DB 299 GRIR0DISHPCGLNM 314

 RESULT 9
 CYOD_PSEPU STANDARD; PRT; 110 AA.
 AC Q9WNR4;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome o ubiquinol oxidase protein cyod.
 GN CYOD.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IH-2000;

RX MEDLINE=99085656; PubMed=9868765;
 RA Hirayama H., Takami H., Inoue A., Horikoshi K.;
 RT "Isolation and characterization of toluene-sensitive mutants from
 RT Pseudomonas putida IH-2000."
 RL FEMS Microbiol. Lett. 169:219-225(1998).
 CC -1- FUNCTION: CYTOCHROME O TERMINAL OXIDASE COMPLEX IS THE COMPONENT
 CC OF THE AEROBIC RESPIRATORY CHAIN THAT PREDOMINATES WHEN CELLS ARE
 CC GROWN AT HIGH AERATION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (By similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB016787; BAA76359.1;
 KM Oxidoreductase; Electron transport; Transmembrane; Inner membrane.
 FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 19 39 POTENTIAL.
 FT DOMAIN 40 46 PERIPLASMIC (POTENTIAL).
 FT TRANSSEM 47 67 POTENTIAL.
 FT DOMAIN 68 78 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 79 99 POTENTIAL.
 FT DOMAIN 100 110 PERIPLASMIC (POTENTIAL).
 SQ SEQUENCE 110 AA; 12376 MM; 074ED0BE85FC0C9 CRC64;

Query Match 32.1%; Score 46.5; DB 1; Length 110;
 Best Local Similarity 40.0%; Pred. No. 4.9;
 Matches 12; Conservative 1; Mismatches 6; Indels 11; Gaps 1;
 OY 4 VIOVISH-----ECRANSMOFLFT 22
 DB 56 VIOVYVHLVYELHMDRSKEORNNWTFLLFT. 85

 RESULT 10
 INX2_SCHAM STANDARD; PRT; 359 AA.
 ID INX2_SCHAM
 AC Q9XYN1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Innexin Inx2 (Innexin-2) (G-Inx2).
 GN INX2.
 OS Schistocerca americana (American grasshopper).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Phryganea; Neoptera; Orthopteroidea; Orthoptera; Cellifera;
 OC Acridoidea; Acrididae; Schistocerca.
 OX NCBI_TaxID=7009;
 RN [1]
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
 RP TISSUE-Embryo, Ventral nerve cord, and Body wall;
 RX MEDLINE=99179234; PubMed=10079517;
 RA Canforina M.D., Sanchez D., Herrera M., Bastiani M.J.;
 RT "Developmental expression and molecular characterization of two gap
 RT junction channel proteins expressed during embryogenesis in the
 RT grasshopper Schistocerca americana."
 RL Dev. Genet. 24:137-150(1999).
 CC -1- FUNCTION: STRUCTURAL COMPONENTS OF THE GAP JUNCTIONS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- TISSUE SPECIFICITY: WIDESPREAD EXPRESSION IN EMBRIO, IN ANTERIOR
 CC AND POSTERIOR ROW OF NEURAL PRECURSORS, MIDLINE PRECURSORS AND IN
 CC EPIHELIAL SHEET OF STOMODEUM.
 CC -1- DEVELOPMENTAL STAGE: EMBRYONIC DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE INNEKIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF115854; AAD29306.1; -
 DR InterPro: IPR000990; Innexin.
 DR Pfam: PF00876; Innexin; 1.
 DR PRINTS: PR01262; INNEXIN.
 KW Gap junction; Transmembrane.
 FT DOMAIN 1 22
 FT TRANSMEM 23 43
 FT DOMAIN 44 109
 FT TRANSMEM 110 130
 FT DOMAIN 131 180
 FT TRANSMEM 181 201
 FT TRANSMEM 202 266
 FT TRANSMEM 267 287
 FT DOMAIN 288 359
 FT SEQUENCE 359 AA; 41445 MW; 9C313EEFDF14F1EE CRC64;

Query Match 31.7%; Score 46; DB 1; Length 359;
 Best Local Similarity 41.2%; Pred. No. 16;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 QY 5 IGTISHECRNMSWQFLF 21
 DB 315 IETISNKCQIGDWFLVX 331

RESULT 11
 ID GLN2_FRAAL STANDARD; PRT; 352 AA.
 AC P20805;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Glutamine synthetase II (EC 6.3.1.2) (Glutamate--ammonia ligase II) (GSI1).
 GN GLNII.
 OS Frankia alni.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Frankineae; Frankiaceae; Frankia.
 OX NCBI_TaxID=1859;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CP11;
 RX MEDLINE=90368593; PubMed=1975584;
 RA Rochefort D.A., Benson D.R.;
 RT "Molecular cloning, sequencing, and expression of the glutamine
 RT synthetase II (glnII) gene from the actinomycete root nodule symbiont
 RT Frankia sp. strain CP11."
 RT J. Bacteriol. 172:5335-5342(1990).
 RL [2]
 RP SEQUENCE OF 1-9 FROM N.A.
 RC STRAIN=CP11;
 RX MEDLINE=93373742; PubMed=8099074;
 RA Hosted T.J., Rochefort D.A., Benson D.R.;
 RT "Close linkage of genes encoding glutamine synthetases I and II in
 RT Frankia alni CP11."
 RT J. Bacteriol. 175:3679-3684(1993).
 CC -1- FUNCTION: ASSIMILATES AMMONIA DURING NITROGEN LIMITATION. MAY ALSO
 CC HAVE A ROLE IN PLANT SYMBIOSIS.
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) -> ADP + phosphate +
 CC L-glutamine.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- INDUCTION: BY NITROGEN STARVATION.
 CC -1- MISCELLANEOUS: TWO FORMS OF GLUTAMINE SYNTHETASE (GSI AND GSI1)
 CC CAN BE FOUND IN THIS NITROGEN FIXING BACTERIA. GSI IS A TYPICAL
 CC PROKARYOTIC GLUTAMINE SYNTHETASE WHEREAS GSI1 IS SIMILAR TO THE

CC ENKARYOTIC ENZYME.
 CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M58415; AAA62803.1; -
 DR EMBL: L10632; -; NOT_ANNOTATED_CDS.
 DR PIR: A36725; AJRKO.
 DR InterPro: IPR001691; GLN_synth.
 DR Pfam: PF00120; gln-synt; 1.
 DR PROSITE: PS00180; GLN_1; 1.
 DR PROSITE: PS00181; GLN_ATP; 1.
 DR Nitrogen fixation; Ligase; Multigene family.
 KW Nitrogen fixation; Ligase; Multigene family.
 SEQUENCE 352 AA; 38665 MW; 2BFCC8BA87C3340 CRC64;

Query Match 31.0%; Score 45; DB 1; Length 352;
 Best Local Similarity 33.3%; Pred. No. 22;
 Matches 8; Conservative 2; Mismatches 14; Indels 0; Gaps 0;
 QY 1 GGGVIGTISHECRNMSWQFLTC 24
 DB 222 GAGAHNTNSTKQIMEGDAIATVC 245

RESULT 12
 ID Y061_MYCGE STANDARD; PRT; 461 AA.
 AC P47307;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MG061.
 GN MG061.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
 RA Nguyen D.T., Uutterback T.R., Sauder D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium."
 RT Science 270:397-403(1995).
 RL [2]
 RP SEQUENCE OF 205-424 FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III,
 RA "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing."
 RT J. Bacteriol. 175:7918-7930(1993).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL: U39685; AAC71278.1; -
DR EMBL: U01705; AAB01017.1; -
DR TIGR: MG061; -
KM Hypothetical protein: Transmembrane: Complete proteome.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 257 277 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 320 340 POTENTIAL.
FT TRANSMEM 356 376 POTENTIAL.
FT TRANSMEM 397 417 POTENTIAL.
SQ SEQUENCE 461 AA; 50972 MW; 319C4906D3E494BB CRC64;

Query Match 31.0%; Score 45; DB 1; Length 461;
Best Local Similarity 30.0%; Pred. No. 28;
Matches 6; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 1 GGGVIGTISHCCRNMSQFL 20
Db 428 GDOLVQEMTKMKLNKQFV 447

RESULT 13
ID Y015_MYCPN STANDARD; PRT; 634 AA.
AC P75084;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ABC transporter ATP-binding protein MG015 homolog
DE (D12.orf634).
GN MN019 OR MP135.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreisch R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF000015; AAB95783.1; -
DR HSSP: P13569; INBD.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_1.
DR SMART: SM00382; AAA_1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1.
DR Hypothetical protein: ATP-binding; Transport; Transmembrane;
KM Complete proteome.
FT TRANSMEM 54 74 POTENTIAL.

```

```

FT TRANSMEM 111 131 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT NP_BIND 430 437 ATP (POTENTIAL).
SQ SEQUENCE 634 AA; 71147 MW; E10CE07E0EB8D406 CRC64;

Query Match 31.0%; Score 45; DB 1; Length 634;
Best Local Similarity 44.4%; Pred. No. 37;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 4 VIGTISHCCRNMSQFL 21
Db 23 VISMLSHNGKPSMKILW 40

RESULT 14
ID FAS2_SCHAM STANDARD; PRT; 898 AA.
AC P22648;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Fasciclin II precursor (FAS II).
DE FAS2.
OS Schistocerca americana (American grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridoidea; Acridoidea; Acrididae; Schistocerca.
OX NCBI_TaxID=7009;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89043938; PubMed=3187519;
RA Harejelson A.L., Goodman C.S.;
RT "Growth cone guidance in insects: fasciclin II is a member of the
RT immunoglobulin superfamily."
RL Science 242:700-708(1988).
RN [2]
RP SEQUENCE OF 423-436.
RX MEDLINE=88276943; PubMed=2839842;
RA Snow P.M., Zinn K., Harejelson A.L., McAllister L., Schilling J.,
RA Bastiani M.J., Makk G., Goodman C.S.;
RT "Characterization and cloning of fasciclin I and fasciclin II
RT glycoproteins in the grasshopper."
RL Proc. Natl. Acad. Sci. U.S.A. 85:5291-5295(1988).
CC -1- FUNCTION: NEURONAL RECOGNITION MOLECULE. INVOLVED IN A PATHWAY
CC RECOGNITION FOR AXONS DURING THE DEVELOPMENT OF NERVE FASCICLES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03789; AAA29810.1; -
DR PIR: A40114; A40114.
DR PIR: B31817; B31817.
DR HSSP: P17948; 10SV.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003962; FNIII_repeat.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00041; fn3; 2.
DR Pfam: PF00047; Ig; 5.

```

DR PRINTS; PR00014; FNTYPEIII.
 DR SMART; SM00060; FN3; 2.
 DR SMART; SM00410; IG-like; 1.
 DR SMART; SM00408; IGC2; 4.
 KW Cell adhesion; Glycoprotein; Repeat; Immunoglobulin domain;
 KM Transmembrane; Signal; Neurogenesis.
 FT SIGNAL 22
 FT CHAIN 1
 FT 23 898
 FT 23 764
 FT 765 782
 FT TRANSMEM 783 898
 FT DOMAIN 41 120
 FT DOMAIN 149 210
 FT DOMAIN 241 307
 FT DOMAIN 336 414
 FT DOMAIN 443 516
 FT DOMAIN 548 611
 FT DOMAIN 662 730
 FT DISULFID 48 113
 FT DISULFID 156 203
 FT DISULFID 248 300
 FT DISULFID 343 407
 FT DISULFID 450 509
 FT CARBOHYD 35 35
 FT CARBOHYD 51 51
 FT CARBOHYD 149 149
 FT CARBOHYD 192 192
 FT CARBOHYD 297 297
 FT CARBOHYD 328 328
 FT CARBOHYD 447 447
 FT CARBOHYD 457 457
 FT CARBOHYD 580 580
 FT CARBOHYD 580 580
 SQ SEQUENCE 898 AA; 99064 MW; 079899EA4F39604AC CRC64;

Query Match 31.0%; Score 45; DB 1; Length 898;
 Best Local Similarity 26.3%; Pred. No. 50;
 Matches 10; Conservative 4; Mismatches 10; Indels 14; Gaps 1;
 QY 2 GGVYIOTISHEC-----RMSWQFLFTCCS 25
 DB 792 GGLIALCHRCSCSAKTDSDAKIALSLYWRFLPYCS 829

RESULT 15
 CHS2_USTMA STANDARD; PRT; 186 AA.
 AC P30599;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE Chitin synthase 2 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
 transferase 2) (Fragment).
 GN CHS2.
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92115692; PubMed=1731323;
 RA Bowen A.R., Chen-Wu J.L., Momany M., Young R., Szaniszló P.J.,
 RA Robbins P.W.;
 RT "Classification of fungal chitin synthases";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:519-523(1992).
 CC -!- FUNCTION: PLAYS A MAJOR ROLE IN CELL WALL BIOGENESIS.
 CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + ((1,4)-(N-acetyl-
 beta-D-glucosaminyl))_(N) = UDP + ((1,4)-(N-acetyl-beta-D-
 glucosaminyl))_(N+1).
 CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
 CC -!- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M82959; AAA34225.1; -.
 DR PIR; A45189; A45189.
 DR InterPro: IPR002923; Chitin_synth.
 DR Pfam; PF01644; Chitin_synth. 1.
 DR Prodom: PD002998; Chitin_synth. 1.
 KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
 KW Multigene family.
 FT NON_TER 1 1
 FT NON_TER 186 186
 FT NON_TER 186 186
 SQ SEQUENCE 186 AA; 20867 MW; D6ADE6CAE826E697 CRC64;

Query Match 30.7%; Score 44.5; DB 1; Length 186;
 Best Local Similarity 47.4%; Pred. No. 15;
 Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;
 QY 2 GGVYIOTISHEC---RMSW 17
 DB 9 GGVYKINIAHLCSPRTSRKWTW 27

Search completed: May 24, 2002, 09:57:43
 Job time: 332 sec

DB 24 GNGVITISHCHMNTWQFLFTCCS 48

RESULT 2

ID 09K381 PRELIMINARY; PRT; 47 AA.

AC 09K381; 01-OCT-2000 (TREMBLrel. 15, Created)

RT 01-OCT-2000 (TREMBLrel. 15, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE RUMA PROTEIN PRECURSOR.

GN RUMA OR RUMAI OR RUMAZ.

OS Ruminococcus gnavus.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OX NCBI_TaxID=33038;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=EI;

RA Gomez-Rodriguez A., Ladire M., Marcellie F., Fons M.;

RT "Tripsin-Mediated Transcriptional Regulation of Genes Involved in Biosynthesis of ruminococcin A, a Lanthibiotic Produced by a Ruminococcus gnavus Strain Isolated from a Human Intestinal Microflora."

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=EI;

RA Gomez-Rodriguez A., Ladire M., Marcellie F., Fons M.;

RT "Characterization and distribution of IS_{Rgn1}, a novel insertion sequence of the IS3 family isolated from Ruminococcus gnavus."

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ276653; CAB93673.1; -

DR EMBL: AJ276653; CAB93672.1; -

DR EMBL: AF320327; AAK73195.1; -

DR EMBL: AF320327; AAK73193.1; -

DR EMBL: AF320327; AAK73194.1; -

KW Signal.

FT SIGNAL.

FT CHAIN 1 23 POTENTIAL.

FT SEQUENCE 47 AA; 5360 MW; 2D70514404D1FE99 CRC64;

QY 1 GCGVITISHCHMNTWQFLFTCC 24

DB 24 GNGVITISHCHMNTWQFLFTCC 47

Query Match 83.4%; Score 121; DB 2; Length 47;

Best Local Similarity 79.2%; Pred. No. 1.9e-12;

Matches 19; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 3

ID 050848 PRELIMINARY; PRT; 47 AA.

AC 050848; 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE VARIACIN PRECURSOR.

OS Micrococcus varians.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Micrococcales; Micrococcaceae; Kocuria.

OX NCBI_TaxID=1272;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=96209245; PubMed=8633879;

RA Primrose D., Rekhif N., Pittert A.C., Suri B., Mollet B.;

RT "Variacin, a new lanthionine-containing bacteriocin produced by Micrococcus varians: comparison to the lactacin 481 of Lactococcus lactis."

RL Appl. Environ. Microbiol. 62:1799-1802(1996).

DR EMBL: X93303; CAA63706.1; -

KW Signal.

FT SIGNAL.

FT CHAIN 1 22 POTENTIAL.

FT SEQUENCE 47 AA; 5009 MW; C7BB4BA3FE0D647A CRC64;

QY 1 GCGVITISHCHMNTWQFLFTCCS 25

DB 23 GCGVITISHCHMNTWQFLFTCCS 47

Query Match 79.3%; Score 115; DB 2; Length 47;

Best Local Similarity 80.0%; Pred. No. 1.7e-11;

Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

RESULT 4

ID 099016 PRELIMINARY; PRT; 48 AA.

AC 099016; 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)

DE BUTYRIVIBRIOCIN.

GN BVIA.

OS Butyrivibrio fibrisolvens.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OC Butyrivibrio.

OX NCBI_TaxID=831;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OR36, NOR37, OR78, AND OR76;

RA Whitford M.F., McPherson M.A., Forster R.J., Teather R.M.;

RT "Determination of the structure and distribution of bacteriocin OR79A gene homologs in Butyrivibrio strains."

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF349672; AAK32700.1; -

DR EMBL: AF349665; AAK32693.1; -

DR EMBL: AF349667; AAK32695.1; -

DR EMBL: AF349669; AAK32697.1; -

FT SEQUENCE 48 AA; 5461 MW; D7D332FD36BC7A9 CRC64;

QY 1 GCGVITISHCHMNTWQFLFTCCS 25

DB 24 GDGVFRTISHCHMNTWQFLFTCCS 48

Query Match 79.3%; Score 115; DB 2; Length 48;

Best Local Similarity 72.0%; Pred. No. 1.8e-11;

Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

RESULT 5

ID 099015 PRELIMINARY; PRT; 48 AA.

AC 099015; 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)

DE BUTYRIVIBRIOCIN.

GN BVIA.

OS Butyrivibrio fibrisolvens.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OC Butyrivibrio.

OX NCBI_TaxID=831;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OB247, GS111, OR77, OR84, OR85, AND OB146;

RA Whitford M.F., McPherson M.A., Forster R.J., Teather R.M.;

RT "Determination of the structure and distribution of bacteriocin OR79A gene homologs in Butyrivibrio strains."

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF349674; AAK32702.1; -

DR	EMBL	AF349664	AAK32692.1	-
DR	EMBL	AF349664	AAK32692.1	-
DR	EMBL	AF349670	AAK32698.1	-
DR	EMBL	AF349671	AAK32699.1	-
DR	EMBL	AF349673	AAK32701.1	-
SQ	SEQUENCE	48 AA	5395 MW	D7D332PDFF06C7A9 CR664

```

QY      1 GGGVIQTISHECRMNSWQFLPTCCS 25
          | | | : | | | | | | | | | |
Db     24 GGDVFRITISHECAMNTWFIPTCCS 48

```

RP SEQUENCE FROM N.A.
RC STRAIN=T8;
RX MEDLINE=94292461; pubmed=8021218;
RA Novak J, Caufield P.W., Miller E.J.;
RT "Isolation and biochemical characterization of a novel tetracycline
KT mutacin from *Streptococcus mutans*.";
J. Bacteriol. 176:4316-4320(1994).

RP SEQUENCE FROM N.A.
RC STRAIN=T8;
RX MEDLINE=98121310; PubMed=9461412;
RA Woodruff W.A., Novak J., Gausfield P.W.;
RT "Sequence analysis of mutA and mutM genes involved in the biosynthesis of the lantibiotic mutacin II in Streptococcus mutans.";
RL Gene 206:37-43(1998).

```

SQ      SEQUENCE      53 AA;  6020 MW;  6C3788E2C9EC6525 CRC64;

```

```
QY      3 GVIQTISHECRMNSWQFLFTCC 24
Db      32 GVVPTVSIECRMNSWQHVFETCC 53
```

DT	ID	RESULT
01-OCT-2000	09KWM4	7
01-OCT-2000	09KWM4	PRELIMINARY;
01-OCT-2000	09KWM4	PRT; 57 AA.
01-OCT-2000	09KWM4	(Tremblrel.. 15, Created)
01-OCT-2000	09KWM4	(Tremblrel.. 15, Last sequence update)

RN
 LI
 RP SEQUENCE FROM N.A.
 RA Sashihara T., Kimura H., Higuchi T., Matsusaki H., Sonomoto K.,
 RA Ishizaki A.;
 RT "Staphylococcus warneri truncated nukm, nukA, orfI genes, and the
 RT downstream region.";
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.

RA Kimura H., Matsusaki H., Sashihara T., Sonomoto K., Ishizaki A.;
RT purification and partial identification of bacteriocin ISR-1, a new
RT antibiotic produced by *Pediococcus* sp. ISR-1";
RL Biosci. Biotechnol. Biochem. 62:2341-2345(1998).
SO EMBL: AB034941; BAA95674.1; -
DR SEQUENCE 57 AA; 6403 MW; 9BF0046242995626 CRC64;

Query Match	72.4%;	Score 105;	DB 2;	Length 57;
Best Local Similarity	73.9%;	Pred. No. 9.1e-10;		
Matches 17; Conservative	4;	Mismatches 2;	Indels 0;	Gaps 0;

```
QY      3 GVIQTISHECRMNSWQFLFTCCS 25
          ||| |:|:| |||:|:| ||| |
Db     35 GVIPTVSHDCMNSFQFVFTCCS 57
```

RESULT	8		
P71449			
ID	P71449	PRELIMINARY;	PRN; 51 AA.
AC	P71449;		
DT	01-FEB-1997	(TREMBLrel. 02,	Created)
DT	01-FEB-1997	(TREMBLrel. 02,	Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19,	Last annotation update)
DE	BACTERIOCIN J46 PRECURSOR.		

0C Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
0C Lactococcus.
0X NCBI_TaxID=1358;

RP	SEQUENCE FROM N.A.	
RC	STRAIN-J46	
RA	Hoot E., Meghroun J., Barrena-Gonzalez C., Peltidemann H.;	
RT	"bacteriocin j46", a new bacteriocin produced by <i>Lactococcus lactis</i>	
RT	subsp. cremoris J46 : isolation and characterization of the protein	
RT	and its gene."	
RL	Anaerobe 2:137-145(1996).	
DR	EMBL: X89503; CNA61674.1; -.	
KM	Signal.	
FT	1	24
Q	SEQUENCE	POTENTIAL.
	51 AA; 5650 MW;	87A3950BB005AFC2F CRC64;

Query Match	70.3%	Score 102;	DB 2;	Length 51;
Best Local Similarity	72.0%;	Pred. No. 2.5e-09;		
Matches 18; Conservative	1;	Mismatches 6;	Indels 0;	Gaps 0;

```
QY      1 GGGVIQTISHECRMNSWQFLFTCCS 25  
        | ||| |||| | : ||||  
Db     27 GSGVITHISHEVIYNWSNWFVTCCS 51
```

RESULT	9		
031051			
ID	031051	PRELIMINARY;	PRT; 51 AA
AC	031051;		

DT 01-JUN-1998 (TREMBlrel. 05, Created)
 DT 01-JUN-1998 (TREMBlrel. 05, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE SCNAL.
 GN SCNAL.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FF22;
 RX MEDLINE=93319301; PubMed=8328813;
 RA Hynes W.L., Ferretti J.D., Taag J.R.;
 RT "Cloning of the gene encoding Streptococcus A-FF22, a novel
 RT lantibiotic produced by Streptococcus pyogenes, and determination of
 RT its nucleotide sequence";
 RL Appl. Environ. Microbiol. 59:1969-1971(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FF22;
 RA McLaughlin R.E., Hynes W.L.;
 RT "Complete sequence of the Streptococcus pyogenes FF22 lantibiotic
 RT (scn) gene cluster";
 RL Submitted (Sep-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF026542; BAB92601.1;
 SQ SEQUENCE 51 AA; 5658 MW; 02FB9411D2C85AB1 CRC64;

Query Match 51.7%; Score 75; DB 2; Length 51;
 Best Local Similarity 60.9%; Pred. No. 66-05; 7; Gaps 0;
 Matches 14; Conservative 2; Mismatches 7; Indels 0;
 OY 2 GGVIOFISHECRMSWQFLFTCC 24
 Db 29 GHGVNTISACRMSWQFLFTCC 51

RESULT 10
 O9YZY8 PRELIMINARY; PRT; 131 AA.
 AC O9YZY8;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis G virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC GBV-C/HGV group.
 OX NCBI_TaxID=45255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HGV-VT61;
 RX MEDLINE=99175307; PubMed=10074558;
 RA Naito H., Win K.M., Abe K.;
 RT "Identification of a novel genotype of Hepatitis G virus in southeast
 RT Asia";
 RL J. Clin. Microbiol. 37:1217-1220(1999).
 DR EMBL; AB013187; BAA32159.1;
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 13358 MW; 01B25DA355A90F6D CRC64;

Query Match 44.8%; Score 65; DB 12; Length 131;
 Best Local Similarity 41.7%; Pred. No. 0.0071;
 Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
 OY 2 GGVIOFISHECRMSWQFLFTCC 25
 Db 13 GGLIAPATACRANGQYFLTNCA 36

RESULT 11

O9WB76 PRELIMINARY; PRT; 2942 AA.
 ID O9WB76;
 AC O9WB76;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE GENOME POLYPROTEIN.
 OS GB virus C variant troglodytes.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC GBV-C/HGV group.
 OX NCBI_TaxID=93986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98365989; PubMed=9700632;
 RA Birkenmeyer L.G., Desai S.M., Muerhoff A.S., Leary T.P., Simons J.N.,
 RA Montes C.C., Mushawar I.K.;
 RT "Isolation of a GB virus-related genome from a chimpanzee";
 RL J. Med. Virol. 56:44-51(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Birkenmeyer L.G., Muerhoff A.S., Leary T.P., Simons J.N., Montes C.C.,
 RA Desai S.M., Mushawar I.K.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF070476; AAD31543.1;
 DR HSSP; P27958; 1HEI.
 DR MEROPS; S29.002; -.
 DR MEROPS; U39.001; -.
 DR InterPro; IPR000923; Copper_blue1.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR002518; HCV_NS2.
 DR InterPro; IPR004109; HCV_NS3.
 DR InterPro; IPR00745; HCV_NS4A.
 DR InterPro; IPR001490; HCV_NS4B.
 DR InterPro; IPR002868; HCV_NS5A.
 DR InterPro; IPR002166; HCV_RdRp.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00998; HCV_RdRp; 1.
 DR PROSITE; PS00196; COPPER_BLUE; UNKNOWN.1.
 KW Nonstructural protein; Polyprotein; RNA-directed RNA polymerase.
 SQ SEQUENCE 2942 AA; 317213 MW; 9BD8A294601DF237 CRC64;

Query Match 44.8%; Score 65; DB 12; Length 2942;
 Best Local Similarity 45.8%; Pred. No. 0.21;
 Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
 OY 2 GGVIOFISHECRMSWQFLFTCC 25
 Db 122 GGLTPASHACRVGSSYFLTNCA 145

RESULT 12
 O56073 PRELIMINARY; PRT; 2970 AA.
 ID O56073;
 AC O56073;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE GENOME POLYPROTEIN.
 OS Hepatitis GB virus A.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC GBV-C/HGV group.
 OX NCBI_TaxID=39112;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97126113; PubMed=8971037;
 RA Leary T.P., Desai S.M., Yamaguchi J., Chalmers M.L., Schlauder G.G.,
 RA Dawson G.J., Mushawar I.K.;
 RT "Species-specific variants of GB virus A in captive monkeys";
 RL J. Virol. 70:9028-9030(1996).

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97437486; PubMed=9992019;
RA Leary T.P., Desai S.M., Erker J.C., Mushahwar I.K.;
RT "The sequence and genomic organization of a GB virus A variant
RT isolated from captive tamarins.";
RL J. Gen. Virol. 78:2307-2313(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98120818; PubMed=9460920;
RA Erker J.C., Desai S.M., Leary T.P., Chalmers M.L., Montes C.C.,
RA Mushahwar I.K.;
RT "Genomic analysis of two GB virus A variants isolated from captive
RT monkeys.";
RL J. Gen. Virol. 79:41-45(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA Erker J.C., Desai S.M., Leary T.P., Chalmers M.L., Montes C.C.,
RA Mushahwar I.K.;
RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF023424; AAC40501.1; -.
DR HSSP; P27958; 1HEI.
DR MEROPS; S29.002; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002518; HCV_NS2.
DR InterPro; IPR004109; HCV_NS3.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR000566; Lipocin_CytrABP.
DR Pfam; PF01387; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; HCV_RDRP; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
KW Nonstructural protein; Polypeptide; RNA-directed RNA polymerase.
SQ SEQUENCE 2970 AA; 322103 MW; 8BAF2FD9DB89F5A3 CRC64;

Query Match 44.8%; Score 65; DB 12; Length 2970;
Best Local Similarity 48.0%; Pred. No. 0.21;
Matches 12; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 1 GGGVIGTISHCCRNMSWQFLFTCCS 25
   |||: | | | | |
Db 14 GGGVLFKSKHRCRGVDRFLSNCCS 38

RESULT 13
Q9YZV9 PRELIMINARY; PRT; 131 AA.
AC Q9YZV9;
ID Q9YZV9;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis G virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC GBV-C/HGV group.
OX NCBI_TaxID=45255;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HGV-CB8;
RA Abe K., Naito H.;
RT "HGV/GBV-C.";
RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB013329; BAA32200.1; -.
FT NON_TER 131
SQ SEQUENCE 131 AA; 13551 MW; 296A8359286E10A0 CRC64;

```

```

Query Match 42.8%; Score 62; DB 12; Length 131;
Best Local Similarity 41.7%; Pred. No. 0.022;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGVIGTISHCCRNMSWQFLFTCCS 25
   |::: | | | | |
Db 13 GALLAPATHACRNGQYFLTNCCS 36

RESULT 14
Q96900 PRELIMINARY; PRT; 37 AA.
ID Q96900;
AC Q96900;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE E1 PROTEIN (FRAGMENT).
OS Hepatitis GB virus C.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC GBV-C/HGV group.
OX NCBI_TaxID=39839;
RN [1]
RP SEQUENCE FROM N.A.
RA Simons J.N.;
RT "Translational initiation in GB viruses A and C: evidence for internal
RT ribosome entry and implications on genome organization.";
RL J. Virol. 70:0-0(1996).
DR EMBL; U62537; AAB09532.1; -.
FT NON_TER 37
SQ SEQUENCE 37 AA; 3846 MW; 666F5B3648F0FED CRC64;

Query Match 40.7%; Score 59; DB 12; Length 37;
Best Local Similarity 37.5%; Pred. No. 0.017;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGVIGTISHCCRNMSWQFLFTCCS 25
   |::: | | | | |
Db 13 GALLAPATHACRANGQYFLTNCCA 36

RESULT 15
Q9W7T5 PRELIMINARY; PRT; 50 AA.
ID Q9W7T5;
AC Q9W7T5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis G virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC GBV-C/HGV group.
OX NCBI_TaxID=45255;
RN [1]
RP SEQUENCE FROM N.A.
RA Kondo Y., Mizokami M., Nakano T., Ueda R., Mukaide M., Hiki J.K.,
RA Kato T., Oyunsuren T., Rukhbkayev R., Cooksley W.;
RT "Genotype of GB virus C/hepatitis G virus by molecular evolutionary
RT analysis.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98154925; PubMed=9495537;
RA Kondo Y., Mizokami M., Nakano T., Kato T., Orito E., Ueda R.,
RA Mukaide M., Hiki J.K., Oyunsuren T., Cooksley W.G.;
RT "Genotype of GB virus C/hepatitis G virus by molecular evolutionary
RT analysis.";
RL Virus Res. 52:221-230(1997).
DR EMBL; AB003536; BAA36186.1; -.
DR EMBL; AB003534; BAA36153.1; -.
KW Polypeptide.
FT NON_TER 50
SQ SEQUENCE 50 AA; 50

```

SO SEQUENCE 50 AA; 5182 MM; C77B440E5F2CE4C6 CRC64;

Query Match 40.7%; Score 59; DB 12; Length 50;

Best Local Similarity 37.5%; Pred. No. 0.024; 11; Indels 0; Gaps 0;

Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 2 GGVYIQTISHECRNMSWQFLFTCCS 25

Db 13 GAILADPATHACRANRANGQYFILNCCA 36

Search completed: May 24, 2002, 09:57:21
Job time: 340 sec

Fri May 24 11:22:22 2002

Meller
09/913763
Seq IDs 143 w/
Interf Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:50:25 ; Search time 206.09 seconds
(without alignments)
10.247 Million cell updates/sec

Title: US-09-913-763-1
Perfect score: 31
Sequence: 1 GGGVIO 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 3502263 seqs, 351980561 residues
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/PCYUS.COMB.pep:*
- 2: /cgn2_6/ptodata/2/paa/US06.COMB.pep:*
- 3: /cgn2_6/ptodata/2/paa/US07.COMB.pep:*
- 4: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
- 5: /cgn2_6/ptodata/2/paa/US081.COMB.pep:*
- 6: /cgn2_6/ptodata/2/paa/US082.COMB.pep:*
- 7: /cgn2_6/ptodata/2/paa/US083.COMB.pep:*
- 8: /cgn2_6/ptodata/2/paa/US084.COMB.pep:*
- 9: /cgn2_6/ptodata/2/paa/US085.COMB.pep:*
- 10: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
- 11: /cgn2_6/ptodata/2/paa/US087.COMB.pep:*
- 12: /cgn2_6/ptodata/2/paa/US088.COMB.pep:*
- 13: /cgn2_6/ptodata/2/paa/US089.COMB.pep:*
- 14: /cgn2_6/ptodata/2/paa/US090.COMB.pep:*
- 15: /cgn2_6/ptodata/2/paa/US091.COMB.pep:*
- 16: /cgn2_6/ptodata/2/paa/US092.COMB.pep:*
- 17: /cgn2_6/ptodata/2/paa/US093.COMB.pep:*
- 18: /cgn2_6/ptodata/2/paa/US094.COMB.pep:*
- 19: /cgn2_6/ptodata/2/paa/US095.COMB.pep:*
- 20: /cgn2_6/ptodata/2/paa/US096.COMB.pep:*
- 21: /cgn2_6/ptodata/2/paa/US097.COMB.pep:*
- 22: /cgn2_6/ptodata/2/paa/US098.COMB.pep:*
- 23: /cgn2_6/ptodata/2/paa/US099.COMB.pep:*
- 24: /cgn2_6/ptodata/2/paa/US100.COMB.pep:*
- 25: /cgn2_6/ptodata/2/paa/US101.COMB.pep:*
- 26: /cgn2_6/ptodata/2/paa/US60.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	31	100.0	6 23 US-09-913-763-1	Sequence 1, Appl1
2	31	100.0	25 23 US-09-913-763-3	Sequence 3, Appl1
3	31	100.0	86 26 US-60-099-493-8	Sequence 8, Appl1
4	31	100.0	152 15 US-09-129-377-6	Sequence 6, Appl1
5	31	100.0	152 21 US-09-777-558-6	Sequence 6, Appl1
6	31	100.0	172 15 US-09-129-377-14	Sequence 14, Appl1
7	31	100.0	172 21 US-09-777-558-14	Sequence 14, Appl1

8	31	100.0	182 18 US-09-417-507-44277	Sequence 44277, A
9	31	100.0	183 15 US-09-129-377-20	Sequence 20, Appl1
10	31	100.0	183 21 US-09-777-558-20	Sequence 20, Appl1
11	31	100.0	269 15 US-09-129-377-17	Sequence 17, Appl1
12	31	100.0	269 21 US-09-777-558-17	Sequence 17, Appl1
13	31	100.0	308 1 PCT-US99-22853B-2214	Sequence 2214, Ap
14	31	100.0	308 19 US-09-595-328C-2312	Sequence 2312, Ap
15	31	100.0	360 26 US-60-140-956-2542	Sequence 2542, Ap
16	31	100.0	372 18 US-09-408-020-40	Sequence 40, Appl1
17	31	100.0	372 24 US-10-027-801-40	Sequence 40, Appl1
18	31	100.0	372 24 US-10-027-806-40	Sequence 40, Appl1
19	31	100.0	372 24 US-10-029-120-40	Sequence 40, Appl1
20	31	100.0	372 24 US-10-034-623-40	Sequence 40, Appl1
21	31	100.0	387 21 US-09-708-427-3877	Sequence 3877, Ap
22	31	100.0	387 26 US-60-150-584-624	Sequence 624, App
23	31	100.0	393 22 US-09-846-590B-14	Sequence 14, Appl1
24	31	100.0	406 18 US-09-413-198-2085	Sequence 2085, Ap
25	31	100.0	408 18 US-09-408-020-74	Sequence 74, Appl1
26	31	100.0	408 24 US-10-027-801-74	Sequence 74, Appl1
27	31	100.0	408 24 US-10-027-806-74	Sequence 74, Appl1
28	31	100.0	408 24 US-10-029-120-74	Sequence 74, Appl1
29	31	100.0	408 24 US-10-034-623-74	Sequence 74, Appl1
30	31	100.0	414 26 US-60-146-055-882	Sequence 882, App
31	31	100.0	467 21 US-09-708-427-3876	Sequence 3876, Ap
32	31	100.0	475 20 US-09-614-150-25836	Sequence 25836, A
33	31	100.0	475 20 US-09-614-150-42822	Sequence 42822, A
34	31	100.0	475 26 US-60-167-324-2133	Sequence 2133, Ap
35	31	100.0	475 26 US-60-173-386-1906	Sequence 1906, Ap
36	31	100.0	475 26 US-60-173-464-21568	Sequence 21568, Ap
37	31	100.0	475 26 US-60-175-871-2131	Sequence 2131, Ap
38	31	100.0	475 26 US-60-191-637-25975	Sequence 25975, A
39	31	100.0	475 26 US-60-191-637-42465	Sequence 42465, A
40	31	100.0	475 26 US-60-191-681-20584	Sequence 20584, A
41	31	100.0	475 26 US-60-191-700-2073	Sequence 2073, Ap
42	31	100.0	476 26 US-60-184-775-1945	Sequence 1945, Ap
43	31	100.0	476 21 US-09-708-427-3875	Sequence 3875, Ap
44	31	100.0	476 21 US-09-734-237B-54	Sequence 54, Appl1
45	31	100.0	487 21 US-09-734-237B-56	Sequence 56, Appl1

ALIGNMENTS

RESULT 1
US-09-913-763-1
: Sequence 1, Application US/09913763
: GENERAL INFORMATION:
: APPLICANT: Tagg, et al.
: TITLE OF INVENTION: Lantibiotic
: FILE REFERENCE: 512585-2001
: CURRENT APPLICATION NUMBER: US/09/913,763
: CURRENT FILING DATE: 2001-08-17
: PRIOR APPLICATION NUMBER: NZ 500261
: PRIOR FILING DATE: 1999-10-12
: NUMBER OF SEQ ID NOS: 7
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 6
: TYPE: PRT
: ORGANISM: Streptococcus salivarius
US-09-913-763-1

Query Match 100.0%; Score 31; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGVIO 6
Db 1 GGGVIO 6
RESULT 2

```
US-09-913-763-3
; Sequence 3, Application US/09913763
; GENERAL INFORMATION:
; APPLICANT: Tagg, et al.
; TITLE OF INVENTION: Lanthibiotic
; FILE REFERENCE: 512585-2001
; CURRENT APPLICATION NUMBER: US/09/913,763
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: NZ 500261
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Streptococcus salivarius
US-09-913-763-3

Query Match          100.0%; Score 31; DB 23; Length 25;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
Db 1 GGGVQ 6

RESULT 3
US-60-099-493-8
; Sequence 8, Application US/60099493
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Morakinyo, Layo
; APPLICANT: Orozco, Emil M.
; TITLE OF INVENTION: AMINO ACID DECARBOXYLASES
; FILE REFERENCE: BB-1237-P1
; CURRENT APPLICATION NUMBER: US/60/099,493
; CURRENT FILING DATE: 1998-09-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 8
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (85)
US-60-099-493-8

Query Match          100.0%; Score 31; DB 26; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
Db 54 GGGVQ 59

RESULT 4
US-09-129-377-6
; Sequence 6, Application US/09129377A
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Heska Corporation
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH AROMATIC AMINO ACID
; TITLE OF INVENTION: DECARBOXYLASE PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: HM-7
; CURRENT APPLICATION NUMBER: US/09/129,377A
; CURRENT FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Dirofilaria immitis
US-09-129-377-6

Query Match          100.0%; Score 31; DB 15; Length 152;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
Db 26 GGGVQ 31

RESULT 5
US-09-777-558-6
; Sequence 6, Application US/09777558
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Heska Corporation
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH AROMATIC AMINO ACID
; TITLE OF INVENTION: DECARBOXYLASE PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: HM-7
; CURRENT APPLICATION NUMBER: US/09/777,558
; CURRENT FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 09/129,377
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Dirofilaria immitis
US-09-777-558-6

Query Match          100.0%; Score 31; DB 21; Length 152;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
Db 26 GGGVQ 31

RESULT 6
US-09-129-377-14
; Sequence 14, Application US/09129377A
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Heska Corporation
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH AROMATIC AMINO ACID
; TITLE OF INVENTION: DECARBOXYLASE PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: HM-7
; CURRENT APPLICATION NUMBER: US/09/129,377A
; CURRENT FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Dirofilaria immitis
US-09-129-377-14

Query Match          100.0%; Score 31; DB 15; Length 172;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 GGGVIO 6
|||||
Db 57 GGGVIO 62

RESULT 7
US-09-777-558-14

; Sequence 14, Application US/09777558
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Heska Corporation
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH AROMATIC AMINO ACID
; TITLE OF INVENTION: DECARBOXYLASE PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: HW-7
; CURRENT APPLICATION NUMBER: US/09/777,558
; CURRENT FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 09/129,377
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Dirofilaria immitis
US-09-777-558-14

Query Match 100.0%; Score 31; DB 21; Length 172;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
|||||
Db 57 GGGVIO 62

RESULT 8
US-09-417-507-44277

; Sequence 44277, Application US/09417507
; GENERAL INFORMATION:
; APPLICANT: KEITH G. WEINSTOCK ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
; FILE REFERENCE: PATH99-10
; CURRENT APPLICATION NUMBER: US/09/417,507
; CURRENT FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 44312
; SEQ ID NO 44277
; LENGTH: 182
; TYPE: PRT
; ORGANISM: A.fumigatus
US-09-417-507-44277

Query Match 100.0%; Score 31; DB 18; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
|||||
Db 3 GGGVIO 8

RESULT 9
US-09-129-377-20

; Sequence 20, Application US/09129377A
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Heska Corporation
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH AROMATIC AMINO ACID
; TITLE OF INVENTION: DECARBOXYLASE PROTEINS, NUCLEIC ACID MOLECULES, AND

; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: HW-7
; CURRENT APPLICATION NUMBER: US/09/129,377A
; CURRENT FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Brugia malayi
US-09-129-377-20

Query Match 100.0%; Score 31; DB 15; Length 183;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
|||||
Db 57 GGGVIO 62

RESULT 10
US-09-777-558-20

; Sequence 20, Application US/09777558
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Heska Corporation
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH AROMATIC AMINO ACID
; TITLE OF INVENTION: DECARBOXYLASE PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: HW-7
; CURRENT APPLICATION NUMBER: US/09/777,558
; CURRENT FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 09/129,377
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Brugia malayi
US-09-777-558-20

Query Match 100.0%; Score 31; DB 21; Length 183;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
|||||
Db 57 GGGVIO 62

RESULT 11
US-09-129-377-17

; Sequence 17, Application US/09129377A
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Heska Corporation
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH AROMATIC AMINO ACID
; TITLE OF INVENTION: DECARBOXYLASE PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: HW-7
; CURRENT APPLICATION NUMBER: US/09/129,377A
; CURRENT FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Dirofilaria immitis
US-09-129-377-17

Query Match 100.0%; Score 31; DB 15; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
|||||
Db 57 GGGVIO 62

RESULT 12
US-09-777-558-17
; Sequence 17, Application US/09777558
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Hesk Corporation
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH AROMATIC AMINO ACID
; TITLE OF INVENTION: DECARBOXYLASE PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: HW-7
; CURRENT APPLICATION NUMBER: US/09/777,558
; CURRENT FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 09/129,377
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 269
; TYPE: PRT
; ORGANISM: *Dirofilaria immitis*
US-09-777-558-17

Query Match 100.0%; Score 31; DB 21; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
|||||
Db 57 GGGVIO 62

RESULT 13
PCT-US99-22853B-2214
; Sequence 2214, Application PC/TUS9922853B
; GENERAL INFORMATION:
; APPLICANT: Ceres, Inc.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-0569F(PC)
; CURRENT APPLICATION NUMBER: PCT/US99/22853B
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 3938
; SOFTWARE: MS Word 97
; SEQ ID NO 2214
; LENGTH: 308
; TYPE: PRT
; ORGANISM: *Arabidopsis thaliana*
; FEATURE:
; OTHER INFORMATION: LOCATION 1..308, Ceres Seq. ID 1692726
; NAME/KEY: UNSURE
; LOCATION: (1)..(308)
; OTHER INFORMATION: any Xaa = any amino acid, unknown or other
PCT-US99-22853B-2214

Query Match 100.0%; Score 31; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
|||||

Db 4 GGGVIO 9

RESULT 14
US-09-595-328C-2312
; Sequence 2312, Application US/09595328C
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptic
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-0946P
; CURRENT APPLICATION NUMBER: US/09/595,328C
; CURRENT FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3107
; SEQ ID NO 2312
; LENGTH: 308
; TYPE: PRT
; ORGANISM: *Arabidopsis thaliana*
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..308
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..308
; OTHER INFORMATION: Ceres Seq. ID 1012242
US-09-595-328C-2312

Query Match 100.0%; Score 31; DB 19; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
|||||
Db 3 GGGVIO 8

RESULT 15
US-60-140-956-2542
; Sequence 2542, Application US/60140956
; GENERAL INFORMATION:
; APPLICANT: Kerlavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
; FILE REFERENCE: C1000041
; CURRENT APPLICATION NUMBER: US/60/140,956
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 2638
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2542
; LENGTH: 360
; TYPE: PRT
; ORGANISM: *Drosophila*
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(360)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-140-956-2542

Query Match 100.0%; Score 31; DB 26; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
|||||
Db 32 GGGVIO 37

Search completed: May 24, 2002, 09:56:11
Job time: 346 sec

Query Match 96.8%; Score 30; DB 1; Length 23;
Best Local Similarity 83.3%; Pred. No. 6;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
||||:|
Db 1 GGGVQ 6

RESULT 3
PCT-US02-10788-58
; Sequence 58, Application PC/TUS0210788
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Antibodies inhibiting autoimmune disease
; FILE REFERENCE: 1361.005W01
; CURRENT APPLICATION NUMBER: PCT/US02/10788
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 09/828,708
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10788-58

Query Match 96.8%; Score 30; DB 1; Length 23;
Best Local Similarity 83.3%; Pred. No. 6;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
||||:|
Db 1 GGGVQ 6

RESULT 4
PCT-US02-10788-59
; Sequence 59, Application PC/TUS0210788
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Antibodies inhibiting autoimmune disease
; FILE REFERENCE: 1361.005W01
; CURRENT APPLICATION NUMBER: PCT/US02/10788
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 09/828,708
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10788-59

Query Match 96.8%; Score 30; DB 1; Length 23;
Best Local Similarity 83.3%; Pred. No. 6;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
||||:|
Db 1 GGGVQ 6

RESULT 5
US-09-644-668A-15
; Sequence 15, Application US/09644668A
; GENERAL INFORMATION:
; APPLICANT: Korman, Alan J.
; APPLICANT: Halk, Edward L.
; APPLICANT: Lonberg, Nils
; APPLICANT: Medarex, Inc.
; TITLE OF INVENTION: Human CTLA-4 Antibodies and Their Uses
; FILE REFERENCE: 014643-010510US
; CURRENT APPLICATION NUMBER: US/09/644,668A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,452
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: heavy chain variable region predicted sequence for
; OTHER INFORMATION: VH 3-30.3 germline
US-09-644-668A-15

Query Match 96.8%; Score 30; DB 5; Length 98;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
||||:|
Db 8 GGGVQ 13

RESULT 6
US-09-644-668A-21
; Sequence 21, Application US/09644668A
; GENERAL INFORMATION:
; APPLICANT: Korman, Alan J.
; APPLICANT: Halk, Edward L.
; APPLICANT: Lonberg, Nils
; APPLICANT: Medarex, Inc.
; TITLE OF INVENTION: Human CTLA-4 Antibodies and Their Uses
; FILE REFERENCE: 014643-010510US
; CURRENT APPLICATION NUMBER: US/09/644,668A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,452
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: heavy chain variable region predicted sequence for
; OTHER INFORMATION: VH 3-33 germline
US-09-644-668A-21

Query Match 96.8%; Score 30; DB 5; Length 98;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
||||:|
Db 8 GGGVQ 13

RESULT 7


```
PCT-US02-10788-9
; Sequence 9, Application PC/TUS0210788
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Antibodies inhibiting autoimmune disease
; FILE REFERENCE: 1361.005W01
; CURRENT APPLICATION NUMBER: PCT/US02/10788
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 09/828,708
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10788-9
```

```
Query Match          96.8%; Score 30; DB 1; Length 100;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGVQ 6
       11111:1
Db      1 GGGVQ 6
```

```
RESULT 8
PCT-US02-10788-10
; Sequence 10, Application PC/TUS0210788
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Antibodies inhibiting autoimmune disease
; FILE REFERENCE: 1361.005W01
; CURRENT APPLICATION NUMBER: PCT/US02/10788
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 09/828,708
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10788-10
```

```
Query Match          96.8%; Score 30; DB 1; Length 101;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGVQ 6
       11111:1
Db      1 GGGVQ 6
```

```
RESULT 9
PCT-US02-10788-8
; Sequence 8, Application PC/TUS0210788
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Ditzel, H.
; APPLICANT: Burton, D.
; APPLICANT: Schaller, M.
; TITLE OF INVENTION: Antibodies inhibiting autoimmune disease
; FILE REFERENCE: 1361.005W01
```

```
; CURRENT APPLICATION NUMBER: PCT/US02/10788
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 09/828,708
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-10788-8
```

```
Query Match          96.8%; Score 30; DB 1; Length 112;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGVQ 6
       11111:1
Db      1 GGGVQ 6
```

```
RESULT 10
PCT-US02-06679-8
; Sequence 8, Application PC/TUS0206679
; GENERAL INFORMATION:
; APPLICANT: Medimmune, Inc
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING INTERFERIN ALPHA-V-BETA-3 ANTAGONIS
; TITLE OF INVENTION: COMBINATION WITH OTHER PROPHYLACTIC OR THERAPEUTIC AGENTS
; FILE REFERENCE: 10271-053-228
; CURRENT APPLICATION NUMBER: PCT/US02/06679
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/273,098
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/316,321
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US02-06679-8
```

```
Query Match          96.8%; Score 30; DB 1; Length 117;
Best Local Similarity 83.3%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGVQ 6
       11111:1
Db      8 GGGVQ 13
```

```
RESULT 11
PCT-US02-11404-78
; Sequence 78, Application PC/TUS0211404
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
; FILE REFERENCE: PF12P8PCT
; CURRENT APPLICATION NUMBER: PCT/US02/11404
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/283,391
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/317,600
; PRIOR FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 118
; TYPE: PRT
```

ORGANISM: Homo sapiens
PCT-US02-11404-78

Query Match
Best Local Similarity 96.8%; Score 30; DB 1; Length 118;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
||||:|
DB 8 GGGVQ 13

RESULT 12
US-09-644-668A-17
Sequence 17, Application US/09644668A
GENERAL INFORMATION:
APPLICANT: Korman, Alan J.
APPLICANT: Halk, Edward L.
APPLICANT: Lonberg, Nils
APPLICANT: Medarex, Inc.
TITLE OF INVENTION: Human CTLA-4 Antibodies and Their Uses
FILE REFERENCE: 014643-010510US
CURRENT APPLICATION NUMBER: US/09/644,668A
CURRENT FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/150,452
PRIOR FILING DATE: 1999-08-24
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 118
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: heavy chain variable region predicted sequence for
US-09-644-668A-17

Query Match
Best Local Similarity 96.8%; Score 30; DB 5; Length 118;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
||||:|
DB 8 GGGVQ 13

RESULT 13
US-09-644-668A-19
Sequence 19, Application US/09644668A
GENERAL INFORMATION:
APPLICANT: Korman, Alan J.
APPLICANT: Halk, Edward L.
APPLICANT: Lonberg, Nils
APPLICANT: Medarex, Inc.
TITLE OF INVENTION: Human CTLA-4 Antibodies and Their Uses
FILE REFERENCE: 014643-010510US
CURRENT APPLICATION NUMBER: US/09/644,668A
CURRENT FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/150,452
PRIOR FILING DATE: 1999-08-24
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 118
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: heavy chain variable region predicted sequence for
US-09-644-668A-19

Query Match
Best Local Similarity 96.8%; Score 30; DB 5; Length 118;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
||||:|
DB 8 GGGVQ 13

RESULT 14
US-10-120-377-78
Sequence 78, Application US/10120377
GENERAL INFORMATION:
APPLICANT: Rosen, Craig, et al.
TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
FILE REFERENCE: PFI12P8
CURRENT APPLICATION NUMBER: US/10/120,377
CURRENT FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/283,391
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/317,600
PRIOR FILING DATE: 2001-09-07
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.1
SEQ ID NO 78
LENGTH: 118
TYPE: PRT
ORGANISM: Homo sapiens
US-10-120-377-78

Query Match
Best Local Similarity 96.8%; Score 30; DB 6; Length 118;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
||||:|
DB 8 GGGVQ 13

RESULT 15
PCT-US02-11404-76
Sequence 76, Application PC/TUS0211404
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Vascular Endothelial Growth Factor-2
FILE REFERENCE: PFI12P8PCT
CURRENT APPLICATION NUMBER: PCT/US02/11404
CURRENT FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/283,391
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/317,600
PRIOR FILING DATE: 2001-09-07
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn version 3.1
SEQ ID NO 76
LENGTH: 119
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-11404-76

Query Match
Best Local Similarity 96.8%; Score 30; DB 1; Length 119;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
||||:|
DB 8 GGGVQ 13

Fri May 24 11:22:23 2002

us-09-913-763-1.rapn

Page 5

Search completed: May 24, 2002, 09:56:32
Job time: 351 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:56:11 ; Search time 206.09 Seconds
(Without alignments)
42.697 Million cell updates/sec

Title: US-09-913-763-3
Perfect score: 145
Sequence: 1 GGGVYQRTISHECRNMSQFLFTCCS 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/2/paa/PCRTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	145	100.0	25	23	US-09-913-763-3
2	126	86.9	51	1	PCRT-US97-19282-8
3	126	86.9	51	1	PCT-US98-14547-12
4	126	86.9	51	1	PCRT-US98-14547-12
5	126	86.9	51	11	US-08-736-334-8
6	126	86.9	51	15	US-09-151-203-8
7	126	86.9	51	18	US-09-462-478A-12

8	126	86.9	51	20	US-09-642-485A-8	Sequence 8, Appl
9	115	79.3	25	10	US-08-693-353-1	Sequence 1, Appl
10	115	79.3	47	1	PCT-US97-19282-10	Sequence 10, Appl
11	115	79.3	47	10	US-08-693-353-3	Sequence 3, Appl
12	115	79.3	47	11	US-08-736-334-10	Sequence 10, Appl
13	115	79.3	47	11	US-09-151-203-10	Sequence 10, Appl
14	115	79.3	47	20	US-09-642-485A-10	Sequence 10, Appl
15	105	72.4	27	15	US-09-151-203-14	Sequence 14, Appl
16	105	72.4	27	20	US-09-642-485A-14	Sequence 14, Appl
17	105	72.4	53	11	PCT-US97-19282-7	Sequence 7, Appl
18	105	72.4	53	11	US-08-736-334-7	Sequence 7, Appl
19	105	72.4	53	15	US-09-151-203-7	Sequence 7, Appl
20	105	72.4	53	20	US-09-642-485A-7	Sequence 7, Appl
21	100	69.0	51	1	PCT-US97-19282-11	Sequence 11, Appl
22	100	69.0	51	1	PCT-US98-14547-13	Sequence 13, Appl
23	100	69.0	51	1	PCT-US98-14547-13	Sequence 13, Appl
24	100	69.0	51	11	US-08-736-334-11	Sequence 11, Appl
25	100	69.0	51	15	US-09-151-203-11	Sequence 11, Appl
26	100	69.0	51	18	US-09-462-478A-13	Sequence 13, Appl
27	100	69.0	51	20	US-09-642-485A-11	Sequence 11, Appl
28	59	40.7	120	7	US-08-389-886A-358	Sequence 358, App
29	59	40.7	120	8	US-08-463-497A-181	Sequence 181, App
30	59	40.7	120	8	US-08-466-033-181	Sequence 181, App
31	59	40.7	131	7	US-08-389-886A-356	Sequence 356, App
32	59	40.7	131	8	US-08-463-497A-179	Sequence 179, App
33	59	40.7	131	8	US-08-466-033-179	Sequence 179, App
34	59	40.7	139	7	US-08-389-886A-354	Sequence 354, App
35	59	40.7	139	8	US-08-463-497A-177	Sequence 177, App
36	59	40.7	139	8	US-08-466-033-177	Sequence 177, App
37	59	40.7	2873	8	US-08-463-497A-15	Sequence 15, Appl
38	59	40.7	2873	8	US-08-466-033-15	Sequence 15, Appl
39	59	40.7	2873	10	US-08-638-911-12	Sequence 12, Appl
40	59	40.7	2910	7	US-08-389-886A-15	Sequence 15, Appl
41	59	40.7	2910	7	US-08-389-886A-15	Sequence 15, Appl
42	59	40.7	2910	8	US-08-463-497A-183	Sequence 183, App
43	59	40.7	2910	8	US-08-466-033-183	Sequence 183, App
44	59	40.7	2910	22	US-09-828-498-2	Sequence 2, Appl
45	59	40.7	2910	26	US-60-195-597-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-913-763-3
; Sequence 3, Application US/09913763
; GENERAL INFORMATION:
; APPLICANT: Tagg, et al.
; TITLE OF INVENTION: Lantibiotic
; FILE REFERENCE: 512585-2001
; CURRENT APPLICATION NUMBER: US/09/913,763
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: NZ 500261
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Streptococcus salivarius
US-09-913-763-3

Query Match 100.0%; Score 145; DB 23;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 25; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;
OY 1 GGGVYQRTISHECRNMSQFLFTCCS 25
Db 1 GGGVYQRTISHECRNMSQFLFTCCS 25
RESULT 2

TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear
US-08-736-334-8

Query Match 86.9%; Score 126; DB 11; Length 51;
Best Local Similarity 84.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVIGTISHCCRMNSMQFLFTCCS 25
1 ||| ||||| |||||:|||||
Db 27 GSGVIHTISHCCRMNSMQFLFTCCS 51

RESULT 6

US-09-151-203-8
Sequence 8, Application US/09151203
GENERAL INFORMATION:
APPLICANT: Caulfield et al.
TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC
TITLE OF INVENTION: COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mcgregor & Adler, LLP
STREET: 8011 Candle Lane
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,203
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Dr. Benjamin A.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5911CIP/D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acid residues
TYPE: amino acids
TOPOLOGY: linear
US-09-151-203-8

Query Match 86.9%; Score 126; DB 15; Length 51;
Best Local Similarity 84.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVIGTISHCCRMNSMQFLFTCCS 25
1 ||| ||||| |||||:|||||
Db 27 GSGVIHTISHCCRMNSMQFLFTCCS 51

RESULT 7

US-09-462-478A-12
Sequence 12, Application US/09462478A
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBBLANCIN LANTIBIOTIC PRODUCED BY BACILLUS SUBTILIS 168

FILE REFERENCE: 8172-8072
CURRENT APPLICATION NUMBER: US/09/462,478A
CURRENT FILING DATE: 2000-04-17
PRIOR APPLICATION NUMBER: US/60/053,035
PRIOR FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 12

LENGTH: 51
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Lacticin 481
US-09-462-478A-12

Query Match 86.9%; Score 126; DB 18; Length 51;
Best Local Similarity 84.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVIGTISHCCRMNSMQFLFTCCS 25
1 ||| ||||| |||||:|||||
Db 27 GSGVIHTISHCCRMNSMQFLFTCCS 51

RESULT 8

US-09-642-485A-8
Sequence 8, Application US/09642485A
GENERAL INFORMATION:
APPLICANT: Caulfield, Page W., Fengxia Chen, Ping
TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC COMPOSITIONS AND METHODS
FILE REFERENCE: UAB-12705/22
CURRENT APPLICATION NUMBER: US/09/642,485A
CURRENT FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: 09/151,203
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 08/736,334
PRIOR FILING DATE: 1996-10-23
PRIOR APPLICATION NUMBER: 08/230,473
PRIOR FILING DATE: 1994-04-20
PRIOR APPLICATION NUMBER: 60/149,812
PRIOR FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patent version 3.1
SEQ ID NO 8
LENGTH: 51
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-642-485A-8

Query Match 86.9%; Score 126; DB 20; Length 51;
Best Local Similarity 84.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 GGGVIGTISHCCRMNSMQFLFTCCS 25
1 ||| ||||| |||||:|||||
Db 27 GSGVIHTISHCCRMNSMQFLFTCCS 51

RESULT 9
US-08-693-353-1
Sequence 1, Application US/08693353
GENERAL INFORMATION:
APPLICANT: MOLLET, BEAT
APPLICANT: PEEB, JOHN
APPLICANT: PRIDMORE, DAVID
APPLICANT: REKHIF, NADJI
APPLICANT: SURL, BRUNO
TITLE OF INVENTION: BACTERIOCIN
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: VOGT & O'DONNELL

```

: STREET: 707 WESTCHESTER AVENUE
: CITY: WHITE PLAINS, NEW YORK
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 10604
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/693,353
: FILING DATE:
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: HARACZ, STEPHEN M.
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (914)328 0055
: TELEFAX: (914)328 0060
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 25 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: MICROCOCCUS VARIANS
: INDIVIDUAL ISOLATE: TWO CLONES CNCM I-1586 and CNCM
: INDIVIDUAL ISOLATE: I-1587
: US-08-693-353-1

Query Match          79.3%; Score 115; DB 10; Length 25;
Best Local Similarity 80.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVITISHCCHMSWQFLFTCCS 25
Db 1 GSGVITPISHCCHMSWQFLFTCCS 25

RESULT 10
PCT-US97-19282-10
: Sequence 10, Application PC/TUS9719282
: GENERAL INFORMATION:
: APPLICANT: Caulfield et al.
: TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC COMPOSITIONS
: TITLE OF INVENTION: AND METHODS
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dr. Benjamin A. Adler
: STREET: 8011 Candle Lane
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Apple
: OPERATING SYSTEM: Macintosh 7.5
: SOFTWARE: Macintosh Word
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US97/19282
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/736,334
: FILING DATE: October 23, 1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Adler, Dr. Benjamin A.
: REGISTRATION NUMBER: 35,423
: REFERENCE/DOCKET NUMBER: D5911CIP
```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 713-777-2321
: TELEFAX: 713-777-6908
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 47 amino acid residues
: TYPE: amino acids
: STRANDEDNESS: single
: TOPOLOGY: linear
: PCT-US97-19282-10

Query Match          79.3%; Score 115; DB 1; Length 47;
Best Local Similarity 80.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVITISHCCHMSWQFLFTCCS 25
Db 23 GSGVITPISHCCHMSWQFLFTCCS 47

RESULT 11
US-08-693-353-3
: Sequence 3, Application US/08693353
: GENERAL INFORMATION:
: APPLICANT: MOLETT, BEAT
: APPLICANT: PEEL, JOHN
: APPLICANT: PRIMORE, DAVID
: APPLICANT: REKHIF, MADJI
: APPLICANT: SURT, BRUNO
: TITLE OF INVENTION: BACTERIOCIN
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: VOGT & O'DONNELL
: STREET: 707 WESTCHESTER AVENUE
: CITY: WHITE PLAINS, NEW YORK
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 10604
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/693,353
: FILING DATE:
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: HARACZ, STEPHEN M.
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (914)328 0055
: TELEFAX: (914)328 0060
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 47 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-693-353-3

Query Match          79.3%; Score 115; DB 10; Length 47;
Best Local Similarity 80.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVITISHCCHMSWQFLFTCCS 25
Db 23 GSGVITPISHCCHMSWQFLFTCCS 47

RESULT 12
US-08-736-334-10
```



```
; Sequence 10, Application US/08736334
; GENERAL INFORMATION:
; APPLICANT: Caufield et al.
; TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Macintosh Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/736,334
; FILING DATE: October 23, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Dr. Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5911CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-777-2321
; TELEFAX: 713-777-6908
; INFORMATION FOR SEQ. ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acid residues
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-736-334-10
;
Query Match 79.3%; Score 115; DB 11; Length 47;
Best Local Similarity 80.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVITISHECRMNSQFLFTCCS 25
   1 111 111111 111:111111
Db 23 GSGVPTISHECHMNSFOVFTCCS 47

RESULT 13
US-09-151-203-10
; Sequence 10, Application US/09151203
; GENERAL INFORMATION:
; APPLICANT: Caufield et al.
; TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,203
; CLASSIFICATION:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
```

```
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Dr. Benjamin A.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5911CIP/D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ. ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acid residues
; TYPE: amino acids
; TOPOLOGY: linear
;
; US-09-151-203-10
;
Query Match 79.3%; Score 115; DB 15; Length 47;
Best Local Similarity 80.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVITISHECRMNSQFLFTCCS 25
   1 111 111111 111:111111
Db 23 GSGVPTISHECHMNSFOVFTCCS 47

RESULT 14
US-09-642-485A-10
; Sequence 10, Application US/09642485A
; GENERAL INFORMATION:
; APPLICANT: Caufield, Page W. Qi, Fengxia Chen, Ping
; TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC COMPOSITIONS AND METHODS
; FILE REFERENCE: UAB-12705/22
; CURRENT APPLICATION NUMBER: US/09/642,485A
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 09/151,203
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 08/736,334
; PRIOR FILING DATE: 1996-10-23
; PRIOR APPLICATION NUMBER: 08/230,473
; PRIOR FILING DATE: 1994-04-20
; PRIOR APPLICATION NUMBER: 60/149,812
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ. ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ. ID NO 10
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Streptococcus mutans
;
; US-09-642-485A-10
;
Query Match 79.3%; Score 115; DB 20; Length 47;
Best Local Similarity 80.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGVITISHECRMNSQFLFTCCS 25
   1 111 111111 111:111111
Db 23 GSGVPTISHECHMNSFOVFTCCS 47

RESULT 15
US-09-151-203-14
; Sequence 14, Application US/09151203
; GENERAL INFORMATION:
; APPLICANT: Caufield et al.
; TITLE OF INVENTION: NOVEL LANTHIONINE ANTIBIOTIC
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
```

```
? COUNTRY: USA
? ZIP: 77071
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: Apple Macintosh
? OPERATING SYSTEM: Macintosh
? SOFTWARE: Microsoft Word for Macintosh
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/151,203
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Adler, Dr. Benjamin A.
? REGISTRATION NUMBER: 35,423
? REFERENCE/DOCKET NUMBER: D5911CIP/D
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (713) 777-2321
? TELEFAX: (713) 777-6908
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 27 amino acid residues
? TYPE: amino acids
? TOPOLOGY: linear
? US-09-151-203-14

Query Match 72.4%; Score 105; DB 15; Length 27;
Best Local Similarity 72.7%; Pred. No. 6.1e-07;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 GVIQITISHECRMSWQFLFTCC 24
Db 6 GVVFVTSYECRMSWQHVFCTC 27
```

Search completed: May 24, 2002, 09:56:12
Job time: 347 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:56:32 ; Search time 15.36 Seconds
(without alignments)
46.075 Million cell updates/sec

Title: US-09-913-763-3
Perfect score: 145
Sequence: 1 GGGVIGRIISHECRMNSMQLFTCCS 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 116914 seqs, 28308587 residues

Total number of hits satisfying chosen parameters: 116914

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PTCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	35.9	23	6	US-10-105-299-5409
2	46	31.7	402	5	US-09-540-209B-7126
3	46	31.7	432	5	US-09-540-209B-7342
4	45	31.0	164	6	US-10-106-698-6461
5	45	31.0	165	6	US-10-106-698-6405
6	44	30.3	886	6	US-10-002-309B-2
7	42	29.0	197	5	US-09-570-856A-18
8	42	29.0	197	5	US-09-570-856B-18
9	42	29.0	197	5	US-09-990-874-9
10	41	28.3	383	5	US-09-540-209B-6815
11	41	28.3	434	1	PCT-US02-09921-1074
12	41	28.3	641	5	US-09-540-209B-9944
13	40.5	27.9	414	6	US-10-121-062-314
14	40.5	27.9	414	6	US-10-063-502-102
15	40.5	27.9	414	6	US-10-063-510-102
16	40.5	27.9	414	6	US-10-063-512-102
17	40.5	27.9	414	6	US-10-063-513-102
18	40.5	27.9	414	6	US-10-063-514-102
19	40.5	27.9	414	6	US-10-063-515-102
20	40.5	27.9	414	6	US-10-063-516-102
21	40.5	27.9	414	6	US-10-063-517-102
22	40.5	27.9	414	6	US-10-063-518-102
23	40.5	27.9	414	6	US-10-063-519-102
24	40.5	27.9	414	6	US-10-063-520-102
25	40.5	27.9	414	6	US-10-063-521-102
26	40.5	27.9	414	6	US-10-063-523-102

27	40.5	27.9	414	6	US-10-063-524-102	Sequence 102, App
28	40.5	27.9	414	6	US-10-063-525-102	Sequence 102, App
29	40.5	27.9	414	6	US-10-063-526-102	Sequence 102, App
30	40.5	27.9	414	6	US-10-063-527-102	Sequence 102, App
31	40.5	27.9	414	6	US-10-063-528-102	Sequence 102, App
32	40.5	27.9	414	6	US-10-063-529-102	Sequence 102, App
33	40.5	27.9	414	6	US-10-063-530-102	Sequence 102, App
34	40.5	27.9	414	6	US-10-063-532-102	Sequence 102, App
35	40.5	27.9	414	6	US-10-063-534-102	Sequence 102, App
36	40.5	27.9	414	6	US-10-063-536-102	Sequence 102, App
37	40.5	27.9	414	6	US-10-063-537-102	Sequence 102, App
38	40.5	27.9	414	6	US-10-063-538-102	Sequence 102, App
39	40.5	27.9	414	6	US-10-063-540-102	Sequence 102, App
40	40.5	27.9	414	6	US-10-063-541-102	Sequence 102, App
41	40.5	27.9	414	6	US-10-063-544-102	Sequence 102, App
42	40.5	27.9	414	6	US-10-063-546-102	Sequence 102, App
43	40.5	27.9	414	6	US-10-063-547-102	Sequence 102, App
44	40.5	27.9	414	6	US-10-063-548-102	Sequence 102, App
45	40.5	27.9	414	6	US-10-063-549-102	Sequence 102, App

ALIGNMENTS

RESULT 1
US-10-105-299-5409
; Sequence 5409, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105, 299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 5409
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-105-299-5409

Query Match 35.9%; Score 52; DB 6; Length 23;
Best Local Similarity 47.1%; Pred. No. 0.23;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 9 SHECRMNSMQLFTCCS 25
DB 3 SGCHRLGRWSSLMWCCS 19
RESULT 2
US-09-540-209B-7126
; Sequence 7126, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Bretton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540, 209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7126
; LENGTH: 402
; TYPE: PRT
; ORGANISM: B. fragilis
; US-09-540-209B-7126

Query Match 31.7%; Score 46; DB 5; Length 402;
Best Local Similarity 39.1%; Pred. No. 23;
Matches 9; Conservative 4; Mismatches 4; Indels 6; Gaps 1;

```

OY 1 GGGVIQTISHCRMNSWQFLFTC 23
    |||:: |: |||:| |
Db 158 GGGILAQIT-----SWQGFIC 174

```

```

RESULT      3
US-09-540-209B-7342
; Sequence 7342, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540, 209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ. ID NOS: 10444
; SEQ. ID NO 7342
; LENGTH: 432
; TYPE: PRT
; ORGANISM: B. fragilis
US-09-540-209B-7342

```

Query Match	31.7%;	Score 46;	DB 5;	Length 432;
Best Local Similarity	44.4%;	Pred. No. 25;		
Matches	8;	Conservative	5;	Mismatches 5;
				Indels 0;
				Gaps 0;

```
QY 1 GGGVIQTISHECRMNSWQ 18
    | ||:|: |:|: |:|
Db 316 GNHVIRTLFHKCKHTSFQ 333
```

```

RESULT      4
US-10-106-698-6461
: Sequence 6461, Application US/10106698
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
: FILE REFERENCE: PA005P1
: CURRENT APPLICATION NUMBER: US/10/106,698
: CURRENT FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: PCT/US00/26524
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: US 60/157,137
: PRIOR FILING DATE: 1999-09-29
: PRIOR APPLICATION NUMBER: US 60/163,280
: PRIOR FILING DATE: 1999-11-03
: NUMBER OF SEQ ID NOS: 8564
: SOFTWARE: PatentIn Ver. 3.0
: SEQ ID NO 6461
: LENGTH: 164
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (1)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (11)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (22)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (47)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (51)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (64)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

```

```

? NAME/KEY MISC_FEATURE
? LOCATION: (65)
? OTHER_INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: MISC_FEATURE
? LOCATION: (79)
? OTHER_INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: MISC_FEATURE
? LOCATION: (92)
? OTHER_INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: MISC_FEATURE
? LOCATION: (107)
? OTHER_INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: MISC_FEATURE
? LOCATION: (110)
? OTHER_INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: MISC_FEATURE
? LOCATION: (123)
? OTHER_INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; AS-10-106-698-6461

```

Query Match	31.0%;	Score 45;	DB 6;	Length 164;
Best Local Similarity	31.6%;	Pred. No. 14;		
Matches	6;	Conservative	5;	Mismatches 8;
				Indels 0;
				Gaps 0;

```
QY      5 IQTISHECRMNSWQFLFTC 23
          :| : | : ||| |
Db     102 VQNVPRACXLCPRWRLFPC 120
```

```

RESULT      5
US-10-106-698-6405
: Sequence 6405, Application US/10106698
: GENERAL INFORMATION:
: APPLICANT: Ruden et al.
: TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypep
: FILE REFERENCE: PA005P1
: CURRENT APPLICATION NUMBER: US/10/106,698
: CURRENT FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: PCT/US00/26524
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: US 60/157,137
: PRIOR FILING DATE: 1999-09-29
: PRIOR APPLICATION NUMBER: US 60/163,280
: PRIOR FILING DATE: 1999-11-03
: NUMBER OF SEQ ID NOS: 8564
: SOFTWARE: PatentIn Ver. 3.0
: SEQ ID NO 6405
: LENGTH: 165
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: MISC_FEATURE
: LOCATION: (1)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (2)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (4)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (5)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: MISC_FEATURE
: LOCATION: (19)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: US-10-106-698-6405

```

Query Match	31.0%;	Score 45;	DB 6;	Length 165;
Best Local Similarity	34.8%;	Pred. NO. 14;		
Matches	8;	Conservative	3;	Mismatches 12; Indels 0; Gaps 0;

OY 13 RMNSWQFLFTCCS 25
DB 231 KMICMEDLYTICS 243

RESULT 11

PCT-US02-09921-1074
Sequence 1074, Application PC/TUS0209921
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: DAFEO, Abel
APPLICANT: JONES, Anissa L.
APPLICANT: TRAN, Alanna-Phung B.
APPLICANT: DAHL, Christopher R.
APPLICANT: GIETZEN, Darryl
APPLICANT: CHINN, Joyce
APPLICANT: DUFOUR, Gerard E.
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: YU, Jimmy Y.
APPLICANT: TUASON, Olivia
APPLICANT: YAP, Pierre E.
APPLICANT: AMSHEY, Stefan R.
APPLICANT: DAUGHERTY, Sean C.
APPLICANT: DAM, Tam C.
APPLICANT: LIU, Tommy F.
APPLICANT: NGUYEN, Duy-Viet An
APPLICANT: KLEFFELD, Yael
APPLICANT: GERSTIN JR., Edward H.
APPLICANT: PERALTA, Careyana H.
APPLICANT: DAVID, Marie H.
APPLICANT: LEWIS, Samantha A.
APPLICANT: CHEN, Alice J.
APPLICANT: PANZER, Scott R.
APPLICANT: HARRIS, Bernard
APPLICANT: FLORES, Vincent
APPLICANT: MARAMBA, Rakesh
APPLICANT: LO, Audrey
APPLICANT: LAN, Ruth Y.
APPLICANT: URASHKA, Michael
FILE REFERENCE: PT-1232 PCT
TITLE OF INVENTION: SECRETORY MOLECULES
CURRENT APPLICATION NUMBER: PCT/US02/09921
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: 60/280,067; 60/280,068; 60/291,280; 60/291,849;
60/291,829; 60/299,428; 60/300,001; 60/299,776
PRIOR FILING DATE: 2001-03-29; 2001-03-29; 2001-05-16; 2001-05-17;
2001-05-17; 2001-06-19; 2001-06-20; 2001-06-20
NUMBER OF SEQ ID NOS: 1146
SOFTWARE: PERL Program
SEQ ID NO 1074
LENGTH: 434
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No: LG:203483.3.orf2:2001JUN22
NAME/KEY: unsure
LOCATION: 72
OTHER INFORMATION: unknown or other
PCT-US02-09921-1074

Query Match 28.3%; Score 41; DB 1; Length 434;
Best Local Similarity 42.9%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

OY 3 GVIQTISHCRNMSWQFLFTC 23
DB 319 GDFQEIYSEC--NTWDCIATC 337

RESULT 12

US-09-540-209B-9944

Sequence 9944, Application US/09540209B
GENERAL INFORMATION:
APPLICANT: Gary L. Bretton
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
FILE REFERENCE: 2709,1001-001
CURRENT APPLICATION NUMBER: US/09/540,209B
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 10444
SEQ ID NO 9944
LENGTH: 641
TYPE: PRT
ORGANISM: B. fragilis
US-09-540-209B-9944

Query Match 28.3%; Score 41; DB 5; Length 641;
Best Local Similarity 42.9%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 4 VICTISHCRNMSW 17
DB 136 LINGVHSCENTSW 149

RESULT 13

US-10-121-062-314
Sequence 314, Application US/10121062
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Collin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430RIC1
CURRENT APPLICATION NUMBER: US/10/121,062
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 314
LENGTH: 414
TYPE: PRT
ORGANISM: Homo sapien
US-10-121-062-314

Query Match 27.9%; Score 40.5; DB 6; Length 414;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

OY 8 ISHECRNMSWQFLFTC 23
DB 121 MNHRTMD-WMPLMNC 135

RESULT 14

Sequence 102, Application US/10063502
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.

US-10-063-502-102

```
; APPLICANT: Gurney,Austin L.
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,502
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 102
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-063-502-102
```

```
Query Match          27.9%; Score 40.5; DB 6; Length 414;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
```

```
QY      8 ISHECRNMSQFLFTC 23
      :| | | | | | |
Db      121 MNHRTTMD-WMFLWNC 135
```

```
RESULT 15
; US-10-063-510-102
; Sequence 102, Application US/10063510
; GENERAL INFORMATION:
; APPLICANT: Baton,Dan L.
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,Christopher J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,510
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 102
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-063-510-102
```

```
Query Match          27.9%; Score 40.5; DB 6; Length 414;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
```

```
QY      8 ISHECRNMSQFLFTC 23
      :| | | | | | |
Db      121 MNHRTTMD-WMFLWNC 135
```

Search completed: May 24, 2002, 09:56:33
Job time: 352 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:47:45 ; Search time 51.74 Seconds
(without alignments)
12.881 Million cell updates/sec

Title: US-09-913-763-1

Perfect score: 31
Sequence: 1 GCGVIO 6

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	6	22	AAB62355
2	31	100.0	37	22	AAB62356
3	31	100.0	149	22	AAM00995
4	31	100.0	152	21	AA45064
5	31	100.0	159	22	AA40086
6	31	100.0	172	21	AA45066
7	31	100.0	183	21	AA45068
8	31	100.0	269	21	AA45067
9	31	100.0	307	21	AA05607
10	31	100.0	307	21	AA52551
11	31	100.0	372	21	AA90931

12	31	100.0	393	22	AAM50241	Wheat tyrosine dec
13	31	100.0	408	21	AA90948	Cenarchaeum symbio
14	31	100.0	475	22	ABB66348	Drosophila melanog
15	31	100.0	475	22	ABB72010	Drosophila melanog
16	31	100.0	486	22	AAG63571	Amino acid sequenc
17	31	100.0	487	22	AAG63572	Synthetic amino ac
18	31	100.0	490	14	AA834133	C.roseus TDC with
19	31	100.0	497	22	AAM50239	Rice tyrosine deca
20	31	100.0	500	11	AA806829	Tryptamine. Catha
21	31	100.0	500	14	AA834133	Catharantus rosen
22	31	100.0	500	19	AA871745	Catharantus rosen
23	31	100.0	510	22	ABB71951	Drosophila melanog
24	31	100.0	510	22	ABB72012	Drosophila melanog
25	31	100.0	618	22	ABB66120	Drosophila melanog
26	31	96.8	6	19	AAW87430	Peptide determined
27	30	96.8	18	19	AAW69105	Neuronal NOS bindi
28	30	96.8	30	17	AA897323	Humanised monoclon
29	30	96.8	30	17	AA897324	Humanised monoclon
30	30	96.8	50	18	AAW13633	Apollipoprotein (a)
31	30	96.8	61	16	AA876975	HSV-neutralising a
32	30	96.8	62	16	AA876976	Anti-Hi112 antibod
33	30	96.8	95	21	AA840127	Anti-Hi112 antibod
34	30	96.8	98	14	AA834283	Human TNF binding
35	30	96.8	98	21	AA840098	Anti-Hi112 antibod
36	30	96.8	98	21	AA840099	Anti-Hi112 antibod
37	30	96.8	98	21	AA840100	Anti-Hi112 antibod
38	30	96.8	98	21	AA840101	Anti-Hi112 antibod
39	30	96.8	98	21	AA840102	Anti-Hi112 antibod
40	30	96.8	98	21	AA840103	Anti-Hi112 antibod
41	30	96.8	98	21	AA840104	Anti-Hi112 antibod
42	30	96.8	98	21	AA840105	Anti-Hi112 antibod
43	30	96.8	98	21	AA840106	Anti-Hi112 antibod
44	30	96.8	98	21	AA840109	Anti-Hi112 antibod
45	30	96.8	98	21	AA840110	Anti-Hi112 antibod

ALIGNMENTS

```

RESULT 1
AAB62355 standard; peptide; 6 AA.
XX
AC AAB62355;
XX
DT 29-JUN-2001 (first entry)
XX
DE S. salivarius antibacterial protein salivarin B N-terminal fragment.
XX
KW Salivarin B; antibacterial; bacteriocidal; Streptococcus; infection;
KW upper respiratory tract; medicament; dairy product; milk; confectionery;
KW bacteriocin-like inhibitory substance; BLIS; N-terminal.
XX
OS Streptococcus salivarius.
XX
PN WO200127143-A1.
XX
PD 19-APR-2001.
XX
PF 12-OCT-2000; 2000WO-N200197.
XX
PR 12-OCT-1999; 99NZ-0500261.
XX
PA (UYOT-) UNIV OTAGO.
XX
PA (BLIS-) BLIS TECHNOLOGIES LTD.
XX
PI Tagg JR, Dierksen KP, Upton M;
XX
DR WPI; 2001-281973/29.
XX
PT Streptococcus salivarius antibacterial protein, known as salivarin B,
PT useful as an antibiotic for treating infections of the upper
PT respiratory tract caused by streptococcal organisms, especially

```

PT Streptococcus pyogenes -
XX
XX Claim 1; Page 27; 29pp; English.
CC The invention provides an antibacterial protein, salivarin B, derived
CC from Streptococcus salivarius. Salivarin B is bacteriocidal with
CC respect to Streptococcus pyogenes. Salivarin B is useful for treating
CC an individual against infections of the upper respiratory tract caused
CC by streptococcal organisms. An antibacterial composition comprising
CC salivarin B or an organism which can express salivarin B may be part
CC of a medicament such as syrup, mouthwash, gargle, toothpaste or mouth
CC spray. The antibacterial protein or the organism expressing the protein
CC may be included in a dairy product based food or drink (milk powder, milk
CC biscuits, milk, yoghurt or cheese, flavored milk) or in confectionery
CC (e.g. a chewing gum). The antibacterial composition may further comprise
CC one or more secondary antibacterial agents such as bacteriocin-like
CC inhibitory substance(s) (BLIS). The present sequence represents the
CC N-terminal fragment of the antibacterial protein, salivarin B.
XX
SQ Sequence 6 AA;

Query Match 100.0%; Score 31; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 6, 4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIQ 6
|||||
Db 1 gggviq 6

RESULT 2
AAB62356 standard; Protein; 37 AA.
XX
XX AAB62356;
XX
XX 29-JUN-2001 (first entry)
XX
XX
XX S. salivarius antibacterial protein salivarin B.
XX
XX Salivarin B: antibacterial; bacteriocidal; Streptococcus; infection;
XX upper respiratory tract; medicament; dairy product; milk; confectionery;
XX bacteriocin-like inhibitory substance; BLIS.
XX
XX Streptococcus salivarius.
XX
XX
XX Key Location/Qualifiers
XX Peptide 1..12
XX FT /note= "Leader sequence"
XX FT 13..37
XX FT /note= "specifically claimed mature protein sequence"
XX
XX MO200127143-A1.
XX
XX 19-APR-2001.
XX
XX 12-OCT-2000; 2000MO-NZ00197.
XX
XX 12-OCT-1999; 99NZ-0500261.
XX
XX (UYOT-) UNIV OTAGO.
XX (BLIS-) BLIS TECHNOLOGIES LTD.
XX
XX Tagg JR, Dierksen KP, Upton M;
XX
XX WPI: 2001-281973/29.
XX N-PSDB; AAF57681.
XX
XX Streptococcus salivarius antibacterial protein, known as salivarin B,
XX useful as an antibiotic for treating infections of the upper
XX respiratory tract caused by streptococcal organisms, especially
XX Streptococcus pyogenes -

XX
XX Claim 3; Fig 2; 29pp; English.
XX
XX
CC The invention provides an antibacterial protein, salivarin B, derived
CC from Streptococcus salivarius. Salivarin B is bacteriocidal with
CC respect to Streptococcus pyogenes. Salivarin B is useful for treating
CC an individual against infections of the upper respiratory tract caused
CC by streptococcal organisms. An antibacterial composition comprising
CC salivarin B or an organism which can express salivarin B may be part
CC of a medicament such as syrup, mouthwash, gargle, toothpaste or mouth
CC spray. The antibacterial protein or the organism expressing the protein
CC may be included in a dairy product based food or drink (milk powder, milk
CC biscuits, milk, yoghurt or cheese, flavored milk) or in confectionery
CC (e.g. a chewing gum). The antibacterial composition may further comprise
CC one or more secondary antibacterial agents such as bacteriocin-like
CC inhibitory substance(s) (BLIS). The present sequence represents the
CC antibacterial protein, salivarin B.
XX
SQ Sequence 37 AA;

Query Match 100.0%; Score 31; DB 22; Length 37;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIQ 6
|||||
Db 13 gggviq 18

RESULT 3
AAM00995 standard; Protein; 149 AA.
XX
XX AAM00995;
XX
XX 01-OCT-2001 (first entry)
XX
XX
XX Human bone marrow protein, SEQ ID NO: 496.
XX
XX
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
XX antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
XX immunosuppressive; gene therapy; cytokine cell proliferation;
XX cell differentiation modulator; immune disorder; cancer;
XX human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX
XX Homo sapiens.
XX
XX WO200153453-A2.
XX
XX 26-JUL-2001.
XX
XX 23-DEC-2000; 2000MO-US34960.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX Ken F, Wertman T, Xu C, Xue AU, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Drmanac RT;
XX
XX WPI: 2001-488707/53.
XX N-PSDB; AAH90114.
XX
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful

PT for treating e.g. cancer and immune deficiency disorders -
 PS Claim 10; Page 641; 648pp; English.
 XX
 CC The present sequence is one of 251 novel human polypeptides encoded
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia). Inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence
 CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.
 CC
 SQ Sequence 149 AA;
 QY 1 GGGVIQ 6
 Db 57 gggviq 62
 Query Match 100.0%; Score 31; DB 22; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 4
 ID AAY45064 standard; Protein; 152 AA.
 XX
 AC AAY45064;
 XX
 DT 31-MAY-2000 (first entry)
 XX
 DE D. immitis aromatic amino acid decarboxylase PDIAD0152.
 XX
 KM Aromatic amino acid decarboxylase; ADC; parasitic helminth; infection;
 KM nematocider; antihelminthic; filarid; prevention; treatment; detection;
 KW diagnosis; drug screening; vaccine.
 XX
 OS Dirofilaria immitis.
 OS
 PN WO200008163-A2.
 XX
 PD 17-FEB-2000.
 PD
 PF 05-AUG-1999; 99WO-US17858.
 PF
 PR 05-AUG-1998; 98US-0129377.
 PR
 PA (HESK-) HESKA CORP.
 PA
 PI Tang L;
 PI
 DR WPI: 2000-205716/18.
 DR N-PSDB; AA250814.
 DR
 XX
 PT New isolated aromatic amino acid decarboxylase enzymes, useful for
 PT developing products for the diagnosis of parasitic helminth infections
 PT
 PS Claim 2; Pages 80-81; 93pp; English.
 PS
 CC The present sequence is a Dirofilaria immitis aromatic amino acid
 CC decarboxylase (ADC) (denoted as PDIAD0152) encoded by DNA designated
 CC nDIAD0456. The ADC enzyme has nematocider and
 CC antihelminthic activities. The ADC DNA, its protein, homologs,
 CC antibodies or inhibitors can inhibit moulting of filarid

CC larvae in an animal, therefore are useful in preventing or
 CC treating infections caused by parasitic helminths. The present sequence
 CC can also be used for detection and diagnosis of parasitic helminth
 CC infections, and for drug screening.
 CC
 SQ Sequence 152 AA;
 QY 1 GGGVIQ 6
 Db 26 gggviq 31
 Query Match 100.0%; Score 31; DB 21; Length 152;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 5
 ID AAM00866 standard; Protein; 159 AA.
 XX
 AC AAM00866;
 XX
 DT 01-OCT-2001 (first entry)
 XX
 DE Human bone marrow protein, SEQ ID NO: 342.
 DE
 XX
 KM Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
 KM antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
 KM immunosuppressive; gene therapy; cytokine cell proliferation;
 KM cell differentiation modulator; immune disorder; infection; cancer;
 KM human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
 XX
 OS Homo sapiens.
 OS
 PN WO200153453-A2.
 PN
 PD 26-JUL-2001.
 PD
 PF 23-DEC-2000; 2000WO-US34960.
 PF
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 PR
 PA (HYSEQ-) HYSEQ INC.
 PA
 PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Wang J, Werhman T, Xu C, Xue Au, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Dirmanac RT;
 PI
 DR WPI: 2001-488707/53.
 DR N-PSDB; AAH69985.
 DR
 XX
 PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders -
 PT
 PS Claim 10; Page 437-438; 648pp; English.
 PS
 CC The present sequence is one of 251 novel human polypeptides encoded
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence

CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.

XX
 SQ Sequence 159 AA;

Query Match 100.0%; Score 31; DB 22; Length 159;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
 |||||
 Db 64 gggvitiq 69

RESULT 6

AAV45066
 ID AAV45066 standard; Protein; 172 AA.

XX
 AC AAV45066;

XX
 DT 31-MAY-2000 (first entry)

XX
 DE D. immitis aromatic amino acid decarboxylase PDIADC172.

XX
 KM Aromatic amino acid decarboxylase; ADC; parasitic helminth; infection;
 KW nematocidal; antihelminthic; filarid; prevention; treatment; detection;
 KW diagnosis; drug screening; vaccine.

XX
 OS Dirofilaria immitis.

XX
 PN WO200008163-A2.

XX
 PD 17-FEB-2000.

XX
 PF 05-AUG-1999; 99WO-US17858.

XX
 PR 05-AUG-1998; 98US-0129377.

XX
 PA (HESK-) HESKA CORP.

XX
 PI Tang L;

XX
 DR WPI; 2000-205716/18.

XX
 DR N-PSDB; AAZ50818.

XX
 PT New isolated aromatic amino acid decarboxylase enzymes, useful for
 PT developing products for the diagnosis of parasitic helminth infections

XX
 PS Claim 2; Page 86; 93pp; English.

XX
 CC The present sequence is a Dirofilaria immitis aromatic amino acid
 CC decarboxylase (ADC) (denoted as PDIADC172) encoded by DNA designated
 CC nDIADC517. The ADC enzyme has nematocidal and
 CC antihelminthic activities. The ADC DNA, its protein, homologs, mimetopes,
 CC antibodies or inhibitors can inhibit moulting of filarid
 CC larvae in an animal, therefore are useful in preventing or
 CC treating infections caused by parasitic helminths. The present sequence
 CC can also be used for detection and diagnosis of parasitic helminth
 CC infections, and for drug screening.

XX
 SQ Sequence 172 AA;

Query Match 100.0%; Score 31; DB 21; Length 172;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6

Db 57 gggvitiq 62
 |||||

RESULT 7

AAV45068
 ID AAV45068 standard; Protein; 183 AA.

XX
 AC AAV45068;

XX
 DT 31-MAY-2000 (first entry)

XX
 DE B. malayi aromatic amino acid decarboxylase PBMADC183.

XX
 KM Aromatic amino acid decarboxylase; ADC; parasitic helminth; infection;
 KW nematocidal; antihelminthic; filarid; prevention; treatment; detection;
 KW diagnosis; drug screening; vaccine.

XX
 OS Brugia malayi.

XX
 PN WO200008163-A2.

XX
 PD 17-FEB-2000.

XX
 PF 05-AUG-1999; 99WO-US17858.

XX
 PR 05-AUG-1998; 98US-0129377.

XX
 PA (HESK-) HESKA CORP.

XX
 PI Tang L;

XX
 DR WPI; 2000-205716/18.

XX
 DR N-PSDB; AAZ50822.

XX
 PT New isolated aromatic amino acid decarboxylase enzymes, useful for
 PT developing products for the diagnosis of parasitic helminth infections

XX
 PS Claim 2; Pages 92-93; 93pp; English.

XX
 CC The present sequence is a Brugia malayi aromatic amino acid
 CC decarboxylase (ADC) (denoted as PBMADC183) encoded by DNA designated
 CC nBMAC549. The ADC enzyme has nematocidal and
 CC antihelminthic activities. The ADC DNA, its protein, homologs, mimetopes,
 CC antibodies or inhibitors can inhibit moulting of filarid
 CC larvae in an animal, therefore are useful in preventing or
 CC treating infections caused by parasitic helminths. The present sequence
 CC can also be used for detection and diagnosis of parasitic helminth
 CC infections, and for drug screening.

XX
 SQ Sequence 183 AA;

Query Match 100.0%; Score 31; DB 21; Length 183;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
 |||||
 Db 57 gggvitiq 62

RESULT 8

AAV45067
 ID AAV45067 standard; Protein; 269 AA.

XX
 AC AAV45067;

XX
 DT 31-MAY-2000 (first entry)

XX
 DE D. immitis aromatic amino acid decarboxylase PDIADC269.

KW Aromatic amino acid decarboxylase; ADC; parasitic helminth; infection;
KM nematode; antihelminthic; filarid; prevention; treatment; detection;
XX diagnosis; drug screening; vaccine.
OS Dirofilaria immitis.
XX WO200008163-A2.
XX
PD 17-FEB-2000.
XX
XX 05-AUG-1999; 99WO-US17858.
XX
XX 05-AUG-1998; 98US-0129377.
XX
XX (HESK-) HESKA CORP.
XX
PI Tang L;
XX
DR WPI: 2000-205716/18.
DR N-PSDB; AA250820.
XX
XX
XX New isolated aromatic amino acid decarboxylase enzymes, useful for
PT developing products for the diagnosis of parasitic helminth infections
PT
PT
XX
XX
PS Claim 2; Pages 89-90; 93pp; English.
XX
XX The present sequence is a Dirofilaria immitis aromatic amino acid
CC decarboxylase (ADC) (denoted as PDIADC269) encoded by DNA designated
CC nD1ADC608. The ADC enzyme has nematocidal and
CC antihelminthic activities. The ADC DNA, its protein, homologs, mimetopes,
CC antibodies or inhibitors can inhibit moulting of filarid
CC larvae in an animal, therefore are useful in preventing or
CC treating infections caused by parasitic helminths. The present sequence
CC can also be used for detection and diagnosis of parasitic helminth
CC infections, and for drug screening.
XX
SQ Sequence 269 AA:

Query Match 100.0%; Score 31; DB 21; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIG 6
Db 57 99gvidq 62

RESULT 9
AAG05607
ID AAG05607 standard; Protein; 307 AA.
XX
XX AAG05607;
AC
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 2075.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR
XX
XX 05-MAR-1999; 99US-0123180.
PR

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135533.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145813.
 PR 27-JUL-1999; 99US-0145919.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147935.
 PR 11-AUG-1999; 99US-0148171.
 PR 12-AUG-1999; 99US-0148319.
 PR 13-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158202.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160789.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 25-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 31; DB 21; Length 307;
 Best Local Similarity 100.0%; Pred. No. 3; 7e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0;

QY 1 GGGVITQ 6
 |||||
 Db 3 gggvtd 8

RESULT 10
 AAG52551
 ID AAG52551 standard; Protein; 307 AA.
 XX AC AAG52551.
 XX 18-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 66816.
 XX KW Protein identification: signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EPI033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.


```

Query Match          100.0%; Score 31; DB 22; Length 393;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GGGVIO 6
        |||||
        51 gggvviq 56

RESULT 13
ID AAY90948
XX AAY90948 standard; Protein; 408 AA.
AC AAY90948;
XX
XX 30-AUG-2000 (first entry)
DE Cenarchaeum symbiosum open reading frame protein sequence SEQ ID NO:74.
XX
XX Cenarchaeum symbiosum; non-thermophilic; crenarchaeote; physiology;
XX characterisation; archae; therapeutic; industrial; laboratory.
XX
XX Cenarchaeum symbiosum.
XX
XX WO200018909-A2.
XX
XX 06-APR-2000.
XX
XX 29-SEP-1999; 99WO-US22752.
XX
XX 29-SEP-1998; 98US-0102294.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Swanson RV, Feldman RA, Schleper C;
XX
XX WPI: 2000-293148/25.
XX
XX N-PSDB: AAA55223.
XX
XX New nucleic acids and proteins isolated from the non-thermophilic
XX crenarchaeote Cenarchaeum symbiosum, useful in characterizing the
XX physiology of these archae and in therapeutic, industrial or laboratory
XX techniques -
XX
XX Claim 26; Page 192-193; 210pp; English.
XX
XX AAA55186 to AAA55226 and AAY90913 to AAY90951 represent nucleic acids
XX and proteins isolated from the non-thermophilic crenarchaeote
XX Cenarchaeum symbiosum. The nucleic acids and proteins identified in
XX the present invention are useful in characterising the physiology of
XX these archae and can be used in therapeutic, industrial or laboratory
XX techniques. AAA55227 to AAA55260 represent promoter sequences from
XX Cenarchaeum symbiosum. AAA55261 to AAA55269 represent PCR primers and
XX probes used in examples from the present invention.
XX
XX Sequence 408 AA;
SQ

```

```

DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 25836.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI: 2001-656860/75.
XX
XX N-PSDB: ABL10451.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure; SEQ ID NO 25836; 21pp + Sequence listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 475 AA;
SQ

```

```

Query Match          100.0%; Score 31; DB 21; Length 408;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 GGGVIO 6
        |||||
        166 gggvviq 171

RESULT 14
ID ABB66348
XX ABB66348 standard; Protein; 475 AA.
XX
XX ABB66348;
XX

```

```

RESULT 15
ID ABB72010
XX ABB72010 standard; Protein; 475 AA.
XX
XX ABB72010;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 42822.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX

```

XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL16113.
 XX
 XX
 PR New isolated nucleic acid detection reagent for detecting 1000 or more
 PR genes from Drosophila and for elucidating cell signalling and cell-cell
 PR interactions -
 XX
 PS Disclosure; SEQ ID NO 42822; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (AB101840-ABL16175) and the encoded proteins
 CC (AB157737-AB172072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 475 AA;

Query Match 100.0%; Score 31; DB 22; Length 475;
 Best Local Similarity 100.0%; Pred No. 5.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGVIG 6
 |||||
 Db 140 gggvig 145

Search completed: May 24, 2002, 09:51:36
 Job time: 231 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:49:35 ; Search time 21.74 Seconds
(without alignments)
6.741 Million cell updates/sec

Title: US-09-913-763-1

Perfect score: 31

Sequence: 1 GGGVQ 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	393	4	US-09-377-557-14
2	31	100.0	497	4	US-09-377-557-10
3	30	96.8	30	2	US-08-318-157B-30
4	30	96.8	30	2	US-08-318-157B-31
5	30	96.8	98	1	US-08-211-202-118
6	30	96.8	116	1	US-08-211-202-141
7	30	96.8	117	1	US-07-942-245-18
8	30	96.8	117	1	US-07-942-245-34
9	30	96.8	117	1	US-07-942-245-36
10	30	96.8	117	2	US-08-652-558-3
11	30	96.8	117	2	US-08-653-558-44
12	30	96.8	117	2	US-08-653-558-45
13	30	96.8	117	2	US-08-653-558-46
14	30	96.8	117	3	US-08-545-809A-115
15	30	96.8	117	3	US-08-545-809A-117
16	30	96.8	119	1	US-08-331-398A-46
17	30	96.8	119	1	US-08-331-398A-47
18	30	96.8	119	1	US-08-207-996-22
19	30	96.8	119	2	US-08-303-569B-11
20	30	96.8	119	2	US-08-303-569B-12
21	30	96.8	119	2	US-08-303-569B-13
22	30	96.8	119	2	US-08-303-569B-14
23	30	96.8	119	2	US-08-303-569B-15
24	30	96.8	119	2	US-08-303-569B-16
25	30	96.8	119	2	US-08-303-569B-17
26	30	96.8	119	2	US-08-303-569B-18
27	30	96.8	119	2	US-08-303-569B-19

28	30	96.8	119	2	US-08-303-569B-20	Sequence 20, Appl
29	30	96.8	119	2	US-08-303-569B-21	Sequence 21, Appl
30	30	96.8	119	2	US-08-303-569B-22	Sequence 22, Appl
31	30	96.8	119	2	US-08-303-569B-23	Sequence 23, Appl
32	30	96.8	119	2	US-08-303-569B-24	Sequence 24, Appl
33	30	96.8	119	2	US-08-859-649-20	Sequence 20, Appl
34	30	96.8	119	2	US-08-859-649-21	Sequence 21, Appl
35	30	96.8	119	2	US-08-859-649-22	Sequence 22, Appl
36	30	96.8	119	2	US-08-859-649-23	Sequence 23, Appl
37	30	96.8	119	2	US-08-859-649-24	Sequence 24, Appl
38	30	96.8	119	2	US-08-859-649-31	Sequence 31, Appl
39	30	96.8	119	2	US-08-318-157B-12	Sequence 12, Appl
40	30	96.8	119	2	US-08-318-157B-13	Sequence 13, Appl
41	30	96.8	119	2	US-08-318-157B-14	Sequence 14, Appl
42	30	96.8	119	2	US-08-318-157B-15	Sequence 15, Appl
43	30	96.8	119	2	US-08-318-157B-17	Sequence 17, Appl
44	30	96.8	119	2	US-08-318-157B-58	Sequence 58, Appl
45	30	96.8	119	2	US-08-070-116A-12	Sequence 12, Appl

ALIGNMENTS

```
RESULT 1
US-09-377-557-14
; Sequence 14, Application US/09377557
; Patent No. 6297055
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; TITLE OF INVENTION: Amino Acid Decarboxylases
; FILE REFERENCE: BB-1237
; CURRENT APPLICATION NUMBER: US/09/377,557
; CURRENT FILING DATE: 1999-08-19
; EARLIER APPLICATION NUMBER: 60/099,493
; EARLIER FILING DATE: September 8, 1998
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Triliticum aestivum
US-09-377-557-14

Query Match      100.0%; Score 31; DB 4; Length 393;
Best Local Similarity 100.0%; Pred No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGGVQ 6
      |||||
Db      51 GGGVQ 56

RESULT 2
US-09-377-557-10
; Sequence 10, Application US/09377557
; Patent No. 6297055
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Layo O.
; APPLICANT: Orozco, Emil M. Jr.
; TITLE OF INVENTION: Amino Acid Decarboxylases
; FILE REFERENCE: BB-1237
; CURRENT APPLICATION NUMBER: US/09/377,557
; CURRENT FILING DATE: 1999-08-19
; EARLIER APPLICATION NUMBER: 60/099,493
; EARLIER FILING DATE: September 8, 1998
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 497
```

TYPE: PRT
ORGANISM: Oryza sativa
US-09-377-557-10

Query Match 100.0%; Score 31; DB 4; Length 497;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
|||||
Db 154 GGGVQ 159

RESULT 3
US-08-318-157B-30
Sequence 30, Application US/08318157B
Patent No. 5874540
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CRA HUMANIZED
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,157B
FILING DATE: 05-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-157B-30

Query Match 96.8%; Score 30; DB 2; Length 30;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
|||||
Db 8 GGGVQ 13

RESULT 4
US-08-318-157B-31
Sequence 31, Application US/08318157B
Patent No. 5874540
GENERAL INFORMATION:
APPLICANT: HANSEN, Hans J.

APPLICANT: ARMOUR, Kathryn L.
TITLE OF INVENTION: CDR-GRAFTED TYPE III ANTI-CRA HUMANIZED
TITLE OF INVENTION: MOUSE MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,157B
FILING DATE: 05-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 18733/464
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-318-157B-31

Query Match 96.8%; Score 30; DB 2; Length 30;
Best Local Similarity 83.3%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
|||||
Db 8 GGGVQ 13

RESULT 5
US-08-211-202-118
Sequence 118, Application US/08211202
Patent No. 5565332
GENERAL INFORMATION:
APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matheus
APPLICANT: BAIER, Michael
APPLICANT: JESPEERS, Laurent Stephane Anne Therese
APPLICANT: WINTER, Gregory Paul
TITLE OF INVENTION: Production of chimeric antibodies - a
TITLE OF INVENTION: Combinatorial approach
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
STREET: 6300 Sears Tower, 233 South Wacker Drive
City: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/211,202
FILING DATE: 23-SEP-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120252.3
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120377.8
FILING DATE: 25-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/31960
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 118:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-211-202-118

Query Match 96.8%; Score 30; DB 1; Length 98;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
1111:1
Db 8 GGGVQ 13

RESULT 6
US-08-211-202-141
Sequence 141, Application US/08211202
Patent No. 5565332
GENERAL INFORMATION:
APPLICANT: HOOGENBOOM, Hendricus Renerus Jacobus Matheus
APPLICANT: BAIER, Michael
APPLICANT: JESPER, Laurent Stephane Anne Therese
APPLICANT: WINTER, Gregory Paul
TITLE OF INVENTION: Production of chimeric antibodies - a
TITLE OF INVENTION: combinatorial approach
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: David W. Clough, Marshall O'Toole Gerstein Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211,202
FILING DATE: 23-SEP-1992
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120252.3
FILING DATE: 23-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9120377.8
FILING DATE: 25-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206318.9
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9206372.6
FILING DATE: 24-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/00883
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: David W. Clough
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28111/31960
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 141:
SEQUENCE CHARACTERISTICS:
LENGTH: 116 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-211-202-141

Query Match 96.8%; Score 30; DB 1; Length 116;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
1111:1
Db 8 GGGVQ 13

RESULT 7
US-07-942-245-18
Sequence 18, Application US/07942245
Patent No. 5639641
GENERAL INFORMATION:
APPLICANT: PEDERSEN, Jan T.
APPLICANT: SEARLE, Stephen M.J.
APPLICANT: REES, Anthony R.
APPLICANT: ROGUSKA, Michael A.
APPLICANT: GUILD, Braydon C.
TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
TITLE OF INVENTION: ANTIBODIES
NUMBER OF SEQUENCES: 522
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: in house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860

TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-18

Query Match 96.8%; Score 30; DB 1; Length 117;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
1111:1
DB 8 GGGVQ 13

RESULT 8
US-07-942-245-34
; Sequence 34, Application US/07942245
; Patent No. 5639641

GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: HP 9000/700 Workstation
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-34

Query Match 96.8%; Score 30; DB 1; Length 117;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
1111:1
DB 8 GGGVQ 13

RESULT 9
US-07-942-245-36
; Sequence 36, Application US/07942245

Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: HP 9000/700 Workstation
; OPERATING SYSTEM: UNIX
; SOFTWARE: In house
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,245
; FILING DATE: 09-SEP-1992
; CLASSIFICATION: 530
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-942-245-36

Query Match 96.8%; Score 30; DB 1; Length 117;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
1111:1
DB 8 GGGVQ 13

RESULT 10
US-08-652-558-3
; Sequence 3, Application US/08652558
; Patent No. 5861155
; GENERAL INFORMATION:
; APPLICANT: LIN, AUGUSTINE YEE-THARN
; TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & WITCOFF
; STREET: 75 STATE STREET, 23RD FLOOR
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,558
; FILING DATE: JUNE 6, 1996
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-3

Query Match 96.8%; Score 30; DB 2; Length 117;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
||||:|
Db 8 GGGVQ 13

RESULT 11
US-08-652-558-44
Sequence 44, Application US/08652558
Patent No. 5861155
GENERAL INFORMATION:
APPLICANT: LIN, AUGUSTINE YEE-THARN
TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,558
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-44

Query Match 96.8%; Score 30; DB 2; Length 117;

Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
||||:|
Db 8 GGGVQ 13

RESULT 12
US-08-652-558-45
Sequence 45, Application US/08652558
Patent No. 5861155
GENERAL INFORMATION:
APPLICANT: LIN, AUGUSTINE YEE-THARN
TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,558
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-45

Query Match 96.8%; Score 30; DB 2; Length 117;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
||||:|
Db 8 GGGVQ 13

RESULT 13
US-08-652-558-46
Sequence 46, Application US/08652558
Patent No. 5861155
GENERAL INFORMATION:
APPLICANT: LIN, AUGUSTINE YEE-THARN
TITLE OF INVENTION: HUMANIZED ANTIBODIES AND USES
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF
STREET: 75 STATE STREET, 23RD FLOOR

CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,558
FILING DATE: JUNE 6, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB94/00387
FILING DATE: NOVEMBER 21, 1994
ATTORNEY/AGENT INFORMATION:
NAME: YANKWICH, LEON R.
REGISTRATION NUMBER: 30,237
REFERENCE/DOCKET NUMBER: 95,497-L
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-652-558-46

Query Match 96.8%; Score 30; DB 2; Length 117;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
||||:|
DB 8 GGGVQ 13

RESULT 14
US-08-545-809A-115
Sequence 115, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
REFERENCE/DOCKET NUMBER: 06501/004001

TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-115

Query Match 96.8%; Score 30; DB 3; Length 117;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
||||:|
DB 27 GGGVQ 32

RESULT 15
US-08-545-809A-117
Sequence 117, Application US/08545809A
Patent No. 6096878
GENERAL INFORMATION:
APPLICANT: Honjo, Tasuku
APPLICANT: Matsuda, Fumihiko
TITLE OF INVENTION: HUMAN IMMUNOGLOBULIN VH GENE
TITLE OF INVENTION: SEGMENTS AND DNA FRAGMENTS CONTAINING THE SAME
NUMBER OF SEQUENCES: 145
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,809A
FILING DATE: 27-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP93/00603
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 06501/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 117:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-545-809A-117

Query Match 96.8%; Score 30; DB 3; Length 117;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVQ 6
||||:|

Db 27 GGGVWQ 32

Search completed: May 24, 2002, 09:52:04
Job time: 149 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:49:50 ; Search time 26.69 Seconds
(without alignments)
21.601 Million cell updates/sec

Title: US-09-913-763-1

Perfect score: 31

Sequence: 1 GGGVIO 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	372	2 T31326	hypothetical prote
2	31	100.0	457	2 J02184	hypothetical 50.4K
3	31	100.0	464	2 A28569	alpha-methylidopa-h
4	31	100.0	479	2 A84588	probable tyrosine
5	31	100.0	480	1 DCHUA	aromatic-L-amino-a
6	31	100.0	480	1 DEGPA	aromatic-L-amino-a
7	31	100.0	480	1 DCRTA	aromatic-L-amino-a
8	31	100.0	485	2 S17848	aromatic-L-amino-a
9	31	100.0	487	2 A43758	aromatic-L-amino-a
10	31	100.0	500	1 DCJAP	aromatic-L-amino-a
11	31	100.0	503	1 DCFD2	aromatic-L-amino-a
12	31	100.0	510	1 DCFPA	aromatic-L-amino-a
13	31	100.0	511	1 DCFED1	aromatic-L-amino-a
14	31	100.0	545	2 T10652	aromatic amino-aci
15	31	100.0	650	2 T23175	hypothetical prote
16	30	96.8	94	2 PLO120	Ig heavy chain V-I
17	30	96.8	96	2 PH0873	Ig heavy chain V-I
18	30	96.8	97	2 S44115	Ig heavy chain V-I
19	30	96.8	98	2 PLO116	Ig heavy chain V-I
20	30	96.8	98	2 S29546	Ig heavy chain V-I
21	30	96.8	98	2 S29543	Ig heavy chain V-I
22	30	96.8	98	2 S54856	Ig heavy chain V-I
23	30	96.8	101	2 J70511	Ig heavy chain V-I
24	30	96.8	113	2 S38490	Ig heavy chain V-I
25	30	96.8	113	2 S57410	Ig heavy chain V-I
26	30	96.8	114	2 S46390	Ig heavy chain V-I
27	30	96.8	114	2 S46392	Ig heavy chain V-I
28	30	96.8	115	2 S36267	Ig heavy chain V-I
29	30	96.8	115	2 S57445	Ig heavy chain V-I

30	30	96.8	118	2 S38491	Ig heavy chain - h
31	30	96.8	118	2 S31116	Ig heavy chain - h
32	30	96.8	119	1 G1HUNT	Ig heavy chain V-I
33	30	96.8	119	1 A1HUBR	Ig heavy chain V-I
34	30	96.8	119	2 S37453	Ig mu chain - huma
35	30	96.8	119	2 F36005	Ig heavy chain V-I
36	30	96.8	119	2 S31112	Ig heavy chain - h
37	30	96.8	120	2 S31117	Ig heavy chain - h
38	30	96.8	120	2 S36278	Ig heavy chain V-I
39	30	96.8	121	1 G1HUNT	Ig heavy chain V-I
40	30	96.8	121	2 S19666	Ig heavy chain V-I
41	30	96.8	121	2 G36005	Ig heavy chain V-I
42	30	96.8	122	2 E36005	Ig heavy chain V-I
43	30	96.8	122	2 S31117	Ig heavy chain V-I
44	30	96.8	122	2 S31119	Ig heavy chain - h
45	30	96.8	122	2 S69910	Ig V-D-U region (K

ALIGNMENTS

RESULT 1
T31326
hypothetical protein - Cenarchaeum symbiosum
C:Species: Cenarchaeum symbiosum
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T31326
R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
J. Bacteriol. 180, 5003-5009, 1998
A:Title: Genomic analysis reveals chromosomal variation in natural populations of the
A:Reference number: 220994; MUID:98422450
A:Accession: T31326
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1372 <SCH>
A:Cross-references: EMBL:AF083072; NID:93599393; PID:93599412; PIDN:MAC62717.1

Query Match 100.0%; Score 31; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
DB 130 GGGVIO 135

RESULT 2
J02184
hypothetical 50.4K protein - apple chlorotic leaf spot virus (isolate P-205)
N:Alternate names: ORF 2 protein
C:Species: apple chlorotic leaf spot virus
C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000
C:Accession: J02184
R:Sato, K.; Yoshikawa, N.; Takahashi, T.
J. Gen. Virol. 74, 1927-1931, 1993
A:Title: Complete nucleotide sequence of the genome of an apple isolate of apple chlo
A:Reference number: J02183; MUID:93389448
A:Accession: J02184
A:Molecule type: mRNA
A:Residues: 1-457 <SAT>
A:Cross-references: GB:D14996; NID:9434059; PIDN:BA03642.1; PID:9453241

Query Match 100.0%; Score 31; DB 2; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
DB 423 GGGVIO 428

RESULT 3
A28569
alpha-methyl-dopa-hypersensitive protein - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 24-Nov-1999
C:Accession: A28569
R:Marsh, J.L.; Ertle, M.P.; Leeds, C.A.
Genetics 114, 453-467, 1986
A:Title: Molecular localization, developmental expression and nucleotide sequence of the A:Reference number: A28569; MUID:87031534
A:Accession: A28569
A:Molecule type: mRNA
A:Residues: 1-464 <MAR>
C:Genetics:
A:Gene: amd
A:Cross-references: FlyBase:FBgn0000075
A:Introns: 60/3
C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase homol

Query Match 100.0%; Score 31; DB 2; Length 464;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
|||||
DB 95 GGGVIO 100

RESULT 4
A84588
probable tyrosine decarboxylase [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: A84588
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Neuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: A84420; MUID:20083487
A:Accession: A84588
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <STO>
A:Cross-references: GB:AE002093; NID:g4512701; PIDN:AA01754.1; GSPDB:GND0139
C:Genetics:
A:Gene: At2g20340
A:Map position: 2
C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase homol

Query Match 100.0%; Score 31; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
|||||
DB 138 GGGVIO 143

RESULT 5
DCH0A
aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) - human
N:Alternate names: DOPA decarboxylase; hydroxytryptophan decarboxylase; tryptophan decarboxylase; Homo sapiens (man)
C:Species: *Homo sapiens* (man)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Mar-2000
C:Accession: A33663; A42205; I54174; I52961; I39381; A45191; B42205; B45191; C42205; C45 R:Ichinose, H.; Kurosawa, Y.; Tltani, K.; Fujita, K.; Nagatsu, T.
Biochem. Biophys. Res. Commun. 164, 1024-1030, 1989
A:Title: Isolation and characterization of a cDNA clone encoding human aromatic L-amino A:Reference number: A33663; MUID:90073624
A:Accession: A33663

A:Molecule type: mRNA
A:Residues: 1-480 <ICH>
A:Cross-references: GB:M6180; NID:g181520; PIDN:AA58437.1; PID:g181521
R:Sunii-Ichinose, C.; Ichinose, H.; Takahashi, E.; Horii, T.; Nagatsu, T.
Biochemistry 31, 2229-2238, 1992
A:Title: Molecular cloning of genomic DNA and chromosomal assignment of the gene for A:Reference number: A42205; MUID:92172843
A:Accession: A42205
A:Molecule type: DNA
A:Residues: 1-6:65-70;103-108;143-149;189-193;236-241;258-263;290-295;313-318;338-343;

A:Experimental source: placenta
A:Note: sequence modified after extraction from NCBI backbone
A:Note: Intron-exon boundaries are shown
R:Scherer, L.J.; McPherson, J.D.; Wasmuth, J.J.; Marsh, J.L.
Genomics 13, 469-471, 1992
A:Title: Human dopa decarboxylase: localization to human chromosome 7p11 and character A:Reference number: I54174; MUID:92307691
A:Accession: I54174
A:Status: preliminary; nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-480 <RES>
A:Cross-references: GB:M88700; NID:g181650; PIDN:AAA20894.1; PID:g181651

R:Craig, S.P.; Thai, A.L.; Weber, M.; Craig, I.W.
Cytogenet. Cell Genet. 61, 114-116, 1992

A:Title: Localisation of the gene for human aromatic L-amino acid decarboxylase (DDC) A:Reference number: I52961; MUID:93009861
A:Accession: I52961
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 68-105 <RE2>

A:Cross-references: GB:S46516; NID:g257537; PIDN:AA02675.1; PID:g257538
R:Le Van Thai, A.; Coste, E.; Allen, J.M.; Palmlter, R.D.; Weber, M.J.
Brain Res. Mol. Brain Res. 17, 227-238, 1993

A:Title: Identification of a neuron-specific promoter of human aromatic L-amino acid A:Reference number: I39381; MUID:93267794
A:Accession: I39381
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-48, 'G', 50-67 <RE3>

A:Cross-references: GB:U05075; NID:g177935; PIDN:AA059432.1; PID:g177937
C:Comment: This enzyme catalyzes the decarboxylation of aromatic amino acids, includi C:Genetics:

A:Gene: GDB:DDC; AACD
A:Cross-references: GDB:127738; OMIM:107930
A:Map position: 7p11-7p11

A:Introns: 67/3; 105/3; 145/3; 190/3; 238/3; 261/1; 293/1; 315/2; 341/1; 347/3; 380/3 C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase ho

C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; pyridoxal phosphate F:2-471/Domain: animal histidine decarboxylase homology <HDC>
F:303/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
|||||
DB 140 GGGVIO 145

RESULT 6
DEGPA
aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) - guinea pig
N:Alternate names: DOPA decarboxylase; hydroxytryptophan decarboxylase; tryptophan de C:Species: *Cavia porcellus* (guinea pig)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Mar-2000
C:Accession: A35710
R:Taketoshi, M.; Horio, Y.; Imamura, I.; Tanaka, T.; Fukui, H.; Wada, H.
Biochem. Biophys. Res. Commun. 170, 1229-1235, 1990
A:Title: Molecular cloning of guinea-pig aromatic-L-amino acid decarboxylase cDNA. A:Reference number: A35710; MUID:90358824
A:Accession: A35710

A:Molecule type: mRNA
A:Residues: 1-480 <TAK>
A:Cross-references: GB:M58049; NID:q191254; PIDN:AAA51530.1; PID:q191255
C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase homol
C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; pyridoxal phosphate
F:2-471/Domain: animal histidine decarboxylase homology <HDC>
F:303/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
|||||
DB 140 GGGVITQ 145

RESULT 7
DCRTA

aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) - rat
N:Alternate names: DOPA decarboxylase; hydroxytryptophan decarboxylase; tryptophan decar
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Mar-2000
C:Accession: A33994; 159415; A47192
R:Tanaka, T.; Horio, Y.; Takekoshi, M.; Imamura, I.; Ando-Yamamoto, M.; Kangawa, K.; Mat
Proc. Natl. Acad. Sci. U.S.A. 86, 8142-8146, 1989
A:Title: Molecular cloning and sequencing of a cDNA of rat dopa decarboxylase: partial
A:Reference number: A33994; MUID:90046754
A:Accession: A33994

A:Molecule type: mRNA
A:Residues: 1-480 <TAN>
A:Cross-references: GB:M27116; NID:q203949; PIDN:AAA41087.1; PID:q203950
R:Rorsman, F.; Husebye, E.S.; Minkvist, O.; Bjork, E.; Karlsson, F.A.; Kampe, O.
Proc. Natl. Acad. Sci. U.S.A. 92, 8626-8629, 1995
A:Title: Aromatic-L-amino-acid decarboxylase, a pyridoxal phosphate-dependent enzyme, is
A:Reference number: 159415; MUID:96004593
A:Accession: 159415

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-197, 'X', 199-475, 'IDFSVTSQ' <RES>
A:Cross-references: EMBL:U31864; NID:q975308; PIDN:AAA85565.1; PID:q975309
A:Experimental source: Insulinoma cells
R:Albert, V.R.; Lee, M.R.; Bolden, A.H.; Wurzbarger, R.J.; Aguanno, A.
Proc. Natl. Acad. Sci. U.S.A. 89, 12053-12057, 1992
A:Title: Distinct promoters direct neuronal and nonneuronal expression of rat aromatic L
A:Reference number: A47192; MUID:93101665

A:Accession: A47192
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-67 <RE2>
A:Cross-references: GB:I03417; NID:q453687; PIDN:AAA99905.1; PID:q453689
C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase homol
C:Keywords: alternative splicing; carbon-carbon lyase; carboxy-lyase; phosphoprotein; py
F:2-471/Domain: animal histidine decarboxylase homology <HDC>
F:303/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
|||||
DB 140 GGGVITQ 145

RESULT 8

SI17848
aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) - pig
N:Alternate names: dopa decarboxylase
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 22-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 07-May-1999

C:Accession: SI17848; S66559
R:Maras, B.; Dominici, P.; Barra, D.; Bossa, F.; Voltattorni, C.B.
Eur. J. Biochem. 201, 385-391, 1991
A:Title: Pig kidney 3,4-dihydroxyphenylalanine (Dopa) decarboxylase. Primary structur
A:Reference number: SI17848; MUID:92037590
A:Accession: SI17848

A:Molecule type: protein
A:Residues: 1-485 <MAR>
R:Moore, P.S.; Dominici, P.; Borri Voltattorni, C.
Biochem. J. 315, 249-256, 1996
A:Title: Cloning and expression of pig kidney dopa decarboxylase: comparison of the n
A:Reference number: S66559; MUID:96207588
A:Accession: S66559

A:Molecule type: mRNA
A:Residues: 207-485, 'S' <MOO>
C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase ho
C:Keywords: acetylated amino end; blocked amino end; carbon-carbon lyase; carboxy-ly
F:2-471/Domain: animal histidine decarboxylase homology <HDC>
F:1/Modified site: acetylated amino end (Met) #status experimental

Query Match 100.0%; Score 31; DB 2; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
|||||
DB 140 GGGVITQ 145

RESULT 9
AA3758

aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 18-Nov-1992 #sequence_revision 18-Nov-1992 #text_change 22-Jun-1999
C:Accession: AA3758
R:Kang, U.U.; Joh, T.H.
Brain Res. Mol. Brain Res. 8, 83-87, 1990
A:Title: Deduced amino acid sequence of bovine aromatic L-amino acid decarboxylase: h
A:Reference number: AA3758; MUID:90339966
A:Accession: AA3758

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-487 <KAN>
A:Cross-references: GB:M74029; NID:q162641; PIDN:AA41615.1; PID:q162642
C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase ho
C:Keywords: carbon-carbon lyase; carboxy-lyase
F:2-471/Domain: animal histidine decarboxylase homology <HDC>

Query Match 100.0%; Score 31; DB 2; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
|||||
DB 140 GGGVITQ 145

RESULT 10
DCUJAP

aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) - Madagascar periwinkle
N:Alternate names: DOPA decarboxylase; hydroxytryptophan decarboxylase; tryptophan de
C:Species: Catharanthus roseus (Madagascar periwinkle)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Mar-2000
C:Accession: A32103; S41726; S25471
R:De Luca, V.; Marneau, C.; Brisson, N.
Proc. Natl. Acad. Sci. U.S.A. 86, 2582-2586, 1989
A:Title: Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxy
A:Reference number: A32103; MUID:89202373
A:Accession: A32103
A:Molecule type: mRNA
A:Residues: 1-500

A:Cross-references: GB:M25151; NID:q167489; PIDN:AAA33109.1; PID:q167490; GB:J04521
 R:Goddijn, O.J.M.; Lohman, F.P.; de Kam, R.J.; Schlipperoot, R.A.; Hoge, J.H.C.
 Mol. Gen. Genet. 242, 217-225, 1994
 A:Title: Nucleotide sequence of the tryptophan decarboxylase gene of *Catharanthus roseus*
 A:Reference number: S41726; MUID:94211212
 A:Accession: S41726
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-500 <GOD>
 A:Cross-references: EMBL:X67662; NID:q18225; PIDN:CAA47898.1; PID:q18226
 C:Comment: This enzyme catalyzes the decarboxylation of aromatic amino acids, including
 C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase homolog
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; pyridoxal phosphate
 F:23-493/Domain: animal histidine decarboxylase homology <HDC>
 F:319/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 1; Length 500;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
 |||||
 Db 159 GGGVIO 164

RESULT 11
 DCEFD2
 aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) 2 - fruit fly (*Drosophila* sp.)
 N:Alternate names: DOPA decarboxylase short isozyme; hydroxytryptophan decarboxylase shc
 C:Species: *Drosophila* sp.
 C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 31-Mar-2000
 R:Evaleth, D.D.; Gietz, R.D.; Spencer, C.A.; Nargang, F.E.; Hodgetts, R.B.; Marsh, J.L.
 EMBO J. 5, 2663-2672, 1986
 A:Title: Sequence and structure of the dopa decarboxylase gene of *Drosophila*: evidence f
 A:Reference number: A91053; MUID:87053836
 A:Accession: B25697
 A:Molecule type: DNA
 A:Residues: 1-503 <EVE>
 A:Cross-references: GB:X04426
 C:Genetics:
 A:Gene: Ddc
 A:Cross-references: FlyBase:FBgn0000422
 A:Introns: 34/2
 C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase homolog
 C:Keywords: alternative splicing; carbon-carbon lyase; carboxy-lyase; phosphoprotein; py
 F:29-496/Domain: animal histidine decarboxylase homology <HDC>
 F:330/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 1; Length 503;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
 |||||
 Db 167 GGGVIO 172

RESULT 12
 DCPFA
 aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) - fruit fly (*Drosophila melanogaster*)
 N:Alternate names: DOPA decarboxylase; hydroxytryptophan decarboxylase; tryptophan decar
 C:Species: *Drosophila melanogaster*
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Jul-2000
 C:Accession: A25709; S01106
 R:Morgan, B.A.; Johnson, W.A.; Hirsch, J.
 EMBO J. 5, 3335-3342, 1986
 A:Title: Regulated splicing produces different forms of dopa decarboxylase in the centra
 A:Reference number: A25709; MUID:87133496
 A:Accession: A25709
 A:Molecule type: DNA

A:Residues: 1-510 <MOR>
 A:Cross-references: GB:X04661; GB:M24111; GB:X16802
 R:Evaleth, D.D.; Marsh, J.L.
 Mol. Gen. Genet. 209, 290-298, 1987
 A:Title: Overlapping transcription units in *Drosophila*: sequence and structure of the
 A:Reference number: S01102; MUID:88038375
 A:Accession: S01106
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 493-510 <EVE>
 A:Cross-references: EMBL:X05991; NID:q7759; PIDN:CAA29409.2; PID:q4455923
 C:Genetics:
 A:Gene: FlyBase:Ddc
 A:Cross-references: FlyBase:FBgn0000422
 A:Introns: 33/3; 62/2
 C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase homolog
 C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; pyridoxal phosphate
 F:37-503/Domain: animal histidine decarboxylase homology <HDC>
 F:337/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
 |||||
 Db 175 GGGVIO 180

RESULT 13
 DCEFD1
 aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) 1 - fruit fly (*Drosophila* sp.)
 N:Alternate names: DOPA decarboxylase long isozyme; hydroxytryptophan decarboxylase 1
 C:Species: *Drosophila* sp.
 C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 31-Mar-2000
 R:Evaleth, D.D.; Gietz, R.D.; Spencer, C.A.; Nargang, F.E.; Hodgetts, R.B.; Marsh, J.
 EMBO J. 5, 2663-2672, 1986
 A:Title: Sequence and structure of the dopa decarboxylase gene of *Drosophila*: evidence
 A:Reference number: A91053; MUID:87053836
 A:Accession: A25697
 A:Molecule type: DNA
 A:Residues: 1-511 <EVE>
 A:Cross-references: GB:X04426
 C:Comment: This enzyme catalyzes the decarboxylation of aromatic amino acids, includi
 C:essential for the central nervous system, and it is necessary in the hypoderm for
 C:Genetics:
 A:Gene: Ddc
 A:Cross-references: FlyBase:FBgn0000422
 A:Introns: 33/3; 62/2
 C:Superfamily: aromatic-L-amino-acid decarboxylase; animal histidine decarboxylase ho
 C:Keywords: alternative splicing; carbon-carbon lyase; carboxy-lyase; phosphoprotein;
 F:337-504/Domain: animal histidine decarboxylase homology <HDC>
 F:338/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 100.0%; Score 31; DB 1; Length 511;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
 |||||
 Db 175 GGGVIO 180

RESULT 14
 T10662
 aromatic amino-acid decarboxylase homolog T5F17.130 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
 C:Accession: T10662
 R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Ban

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16533
 A:Accession: T10662
 A:Molecule type: DNA
 A:Residues: 1-545 <BEV>
 A:Cross-references: EMBL:AL049917; GSPDB:GN00062; ATSP:T5F17.130
 A:Experimental source: cultivar Columbia; BAC clone T5F17
 C:Genetics:
 A:Gene: ATSP:T5F17.130
 A:Map position: 4
 A:Introns: 97/3; 122/1; 199/1; 253/3; 300/3; 320/3; 375/3; 394/3; 416/2; 457/3; 503/3
 C:Superfamily: aromatic-L-aminic acid decarboxylase; animal histidine decarboxylase homol
 F:01-336/Domain: animal histidine decarboxylase homology <HDC>

Query Match 100.0%; Score 31; DB 2; Length 545;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIQ 6
 |||||
 Db 201 GGGVIQ 206

RESULT 15

T23175
 hypothetical protein K01C8.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
 C:Accession: T23175
 R:Stms, M.
 submitted to the EMBL Data Library, April 1995
 A:Reference number: Z19702
 A:Accession: T23175
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-650 <MIT>
 A:Cross-references: EMBL:Z49068; PIDN:CAA8862.1; GSPDB:GN00020; CESP:K01C8.3
 A:Experimental source: clone K01C8
 C:Genetics:
 A:Gene: CESP:K01C8.3
 A:Map position: 2
 A:Introns: 31/1; 74/1; 181/3; 392/2; 416/1; 489/3; 551/1
 C:Superfamily: animal histidine decarboxylase homology
 F:78-546/Domain: animal histidine decarboxylase homology <HDC>

Query Match 100.0%; Score 31; DB 2; Length 650;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIQ 6
 |||||
 Db 217 GGGVIQ 222

Search completed: May 24, 2002, 09:52:38
 Job time: 168 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:52:11 ; Search time 13.51 Seconds
(without alignments)
17.196 Million cell updates/sec

Title: US-09-913-763-1

Perfect score: 31

Sequence: 1 GGGVITQ 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	403	1 DDC_DROLE	096571 drosophila
2	31	100.0	437	1 DDC_DROSI	096567 drosophila
3	31	100.0	457	1 V51K_ACLISA	P54692 apple chlor
4	31	100.0	480	1 DDC_CAVPO	P22781 cavia porce
5	31	100.0	480	1 DDC_HUMAN	P20711 homo sapien
6	31	100.0	480	1 DDC_MOUSE	088533 mus musculu
7	31	100.0	480	1 DDC_RAT	P14173 rattus norv
8	31	100.0	486	1 DDC_PIG	P80041 sus scrofa
9	31	100.0	487	1 DDC_BOVIN	P27718 bos taurus
10	31	100.0	500	1 DDC_CATRO	P17770 catharanthu
11	31	100.0	508	1 DDC_MANSE	P48861 manduca sex
12	31	100.0	510	1 DDC_DROME	P05031 drosophila
13	31	100.0	510	1 L2AM_DROME	P18486 drosophila
14	30	96.8	119	1 HV3I_HUMAN	P01770 homo sapien
15	30	96.8	119	1 HV3I_HUMAN	P01773 homo sapien
16	30	96.8	121	1 HV3L_HUMAN	P01773 homo sapien
17	30	96.8	126	1 HV3K_HUMAN	P01772 homo sapien
18	30	96.8	342	1 FINO_ECOLI	P18809 escherichia
19	30	96.8	367	1 TRMU_NEIMA	093139 neisseria m
20	30	96.8	367	1 TRMU_NEIMA	093136 neisseria m
21	30	96.8	371	1 DUT_HSV1	P10234 herpes simp
22	30	96.8	758	1 PHSA_SALTY	P37600 salmonella
23	30	96.8	808	1 PLD_PIMBR	004883 pimplinella
24	30	96.8	4725	1 DYHC_DICDI	P34036 dictyostell
25	29	93.5	65	1 YSY6_YEAST	P38374 saccharomyc
26	29	93.5	360	1 RTCA_AQUAE	066884 aquifex aeo
27	29	93.5	452	1 TYDI_PETCR	006085 petroselinu
28	29	93.5	508	1 TYDI_PETCR	006088 petroselinu
29	29	93.5	514	1 TYD2_PETCR	006086 petroselinu
30	29	93.5	516	1 TYD3_PETCR	006087 petroselinu
31	29	93.5	518	1 TYDI_PAPSO	P54768 papaver som
32	29	93.5	523	1 TYD5_PAPSO	P54771 papaver som
33	29	93.5	531	1 TYD2_PAPSO	P54769 papaver som

34	29	93.5	618	1 SL55_RAT	063008 rattus norv
35	29	93.5	637	1 TS03_ABIGR	022340 abies grand
36	29	93.5	643	1 SL55_HUMAN	092910 homo sapien
37	29	93.5	656	1 DCHS_RAT	P16453 rattus norv
38	29	93.5	662	1 DCHS_HUMAN	P19113 homo sapien
39	29	93.5	662	1 DCHS_MOUSE	P22738 mus musculu
40	29	93.5	847	1 DCHS_DROME	005733 drosophila
41	28	90.3	115	1 HV3F_HUMAN	P01767 homo sapien
42	28	90.3	116	1 HV05_CARAU	P19181 carassius a
43	28	90.3	122	1 HV3G_HUMAN	P01768 homo sapien
44	28	90.3	238	1 RU2_SUTSO	09uxa5 sulfolobus
45	28	90.3	299	1 SCRK_BACST	005510 bacillus su

ALIGNMENTS

RESULT	ID	DDC_DROLE	STANDARD:	PRT:	403 AA.
DDC_DROLE	AC	096571;			
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DT	01-MAR-2002	(Rel. 41, Last annotation update)			
DE	Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA decarboxylase) (DDC) (Fragment).				
GN	DDC.				
OS	Drosophila lebanonensis (Fruit fly) (Scaptodrosophila lebanonensis).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Scaptodrosophilidae.				
OX	NCBI_TaxID=7225;				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BEIRUT;				
RX	MEDLINE=99250256; PubMed=10231575;				
RA	Talarekxov A., Saez A.G., Ayala F.J.;				
RT	"A compact gene cluster in Drosophila: the unrelated Cs gene is RT compressed between duplicated amd and Ddc.";				
RL	Gene 231:111-120(1999).				
CC	-1- CATALYTIC ACTIVITY: L-TRYPTOPHAN - TRYPTAMINE + CO(2) (ALSO ACTS ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLMALANINE (DOPPA)).				
CC	-1- COFACTOR: PYRIDOXAL PHOSPHATE.				
CC	-1- SUBUNIT: HOMODIMER (BY SIMILARITY).				
CC	-1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; AF091329; AAC67585.1; -				
DR	FLYBASE; FBgn025671; Dileb\ddc.				
DR	InterPro; IPR002129; Pyridoxal.dec.				
DR	Pfam; PF00282; pyridoxal_dec; 1.				
DR	PRINTS; PR00800; YHDCRBOXLASE.				
KW	LYASE; Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate.				
FT	NON_TER				
FT	BINDING				
FT	SEQUENCE	403 AA; 45133 MW; 558EB9B284D5D283 CAC64;			
QY	1 GGGVITQ 6	100.0%; Score 31; DB 1; Length 403;			
DB	66 GGGVITQ 71	Best Local Similarity 100.0%; Pred. No. 54;			
		Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

```

RESULT 2
DDC_DROSI STANDARD; PRT; 437 AA.
AC 096567;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
DE decarboxylase) (DDC) (Fragment).
GN DDC.
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ST. LUCIA;
RX MEDLINE=99250256; PubMed=10231575;
RA Tataronov A., Saez A.G., Ayala F.J.;
RT "A compact gene cluster in Drosophila: the unrelated Cs gene is
RT compressed between duplicated amd and ddc.";
RL Gene 231:111-120(1999).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLALANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF091327; AAC67580.1; -.
DR FlyBase: FBgn0016334; DslnDdc.
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec.1.
DR PRINTS: PR00800; YHDCRBOXYLASE.
KW lysase; Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate.
FT NON TER 1 264
FT BINDING 1 264 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 437 AA; 49058 MW; 92BF87BF39CCBC33 CRC64;

```

```

RX MEDLINE=93389448; PubMed=8376968;
RA Sato K., Yoshikawa N., Takahashi T.;
RT "Complete nucleotide sequence of the genome of an apple isolate of
RT apple chlorotic leaf spot virus.";
RL J. Gen. Virol. 74:1927-1931(1993).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S35.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14996; BAA03642.1; -.
DR MEROPS: S35.001; -.
DR InterPro: IPR001815; Capilloptase.
DR Pfam: PF02103; Peptidase_S35; 1.
DR PRINTS: PR00995; CAPILLOPTASE.
KW Hydrolase; Serine protease.
SQ SEQUENCE 457 AA; 50454 MW; A8986907688B3A58 CRC64;

```

Query Match 100.0%; Score 31; DB 1; Length 457;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 GGGVITQ 6
    |||||
Db 423 GGGVITQ 428

```

```

RESULT 4
DDC_CAVPO STANDARD; PRT; 480 AA.
AC P22781;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
DE decarboxylase) (DDC).
GN DDC.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90358824; PubMed=2390088;
RA Takeuchi M., Horio Y., Imamura I., Tanaka T., Fukui H., Wada H.;
RT "Molecular cloning of guinea-pig aromatic-L-amino acid decarboxylase
RT cDNA.";
RL Biochem. Biophys. Res. Commun. 170:1229-1235(1990).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLALANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: CATECHOLAMINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M58049; AAA51530.1; -.
DR PIR: A35710; DECPA.
DR InterPro: IPR002129; Pyridoxal_dec.

```

```

DR Pfam: PF00282; pyridoxal_dec. 1.
DR PRINTS: PR00800; YHDCRBOXYLASE.
DR PROSITE: PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase: Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate;
Repeat.
FT BINDING 303 303 PYRIDOXAL PHOSPHATE.
FT DOMAIN 58 178 2 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 58 115 1.
FT REPEAT 118 178 2.
SQ SEQUENCE 480 AA; 54150 MW; 46AB0649DB20F5A4 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
Db 140 GGGVIO 145

RESULT 5
DDC_HUMAN STANDARD; PRT; 480 AA.
AC P20711; Q16723;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
DE decarboxylase) (DDC).
GN DDC OR AADC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Pneurochromatoma;
RX MEDLINE=90073624; PubMed=2590185;
RA Ichinose H., Kurosawa Y., Tiltan K., Fujita K., Nagatsu T.;
RA "Isolation and characterization of a cDNA clone encoding human
RT aromatic L-amino acid decarboxylase."
RL Biochem. Biophys. Res. Commun. 164:1024-1030(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92307691; PubMed=1612608;
RA Scherer L.J., McPherson J.D., Wasmuth J.J., Marsh L.J.;
RT "Human dopa decarboxylase: localization to human chromosome 7p11 and
RT characterization of hepatic cDNAs."
RL Genomics 13:469-471(1992).
RN [3]
RP SEQUENCE FROM N.A. AND DISEASE.
RX MEDLINE=92172843; PubMed=1540578;
RA Sumi-Ichinose C., Ichinose H., Takahashi E., Hori T., Nagatsu T.;
RT "Molecular cloning of genomic DNA and chromosomal assignment of the
RT gene for human aromatic L-amino acid decarboxylase, the enzyme for
RT catecholamine and serotonin biosynthesis."
RL Biochemistry 31:2229-2238(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-Lung.
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-67 FROM N.A.
RX MEDLINE=93287794; PubMed=8510497;
RA Van Thai A., Coste E., Allen J.M., Palmiter R.D., Weber M.J.;
RT "Identification of a neuron-specific promoter of human aromatic
RT L-amino acid decarboxylase gene."
RL Brain Res. Mol. Brain Res. 17:227-238(1993).
RN [6]
RP SEQUENCE OF 68-105 FROM N.A.
RX MEDLINE=93009861; PubMed=1395716;

```

```

RA Craig S.P., Thai A.L., Weber M., Craig I.W.;
RT "Localisation of the gene for human aromatic L-amino acid
RT decarboxylase (DDC) to chromosome 7p13-->p11 by in situ
RT hybridisation."
RL Cytogenet. Cell Genet. 61:114-116(1992).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLALANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: CATECHOLAMINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- DISEASE: DEFECTS IN DDC LEAD TO SEVERE HYPOTONIA AND DEVELOPMENTAL
CC DELAY.
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M76180; AAA58437.1; -
DR EMBL: M87700; AAA20894.1; -
DR EMBL: M84592; AAD40482.1; -
DR EMBL: M84600; AAD40482.1; JOINED.
DR EMBL: M84593; AAD40482.1; JOINED.
DR EMBL: M84594; AAD40482.1; JOINED.
DR EMBL: M84596; AAD40482.1; JOINED.
DR EMBL: M84597; AAD40482.1; JOINED.
DR EMBL: M84595; AAD40482.1; JOINED.
DR EMBL: M84598; AAD40482.1; JOINED.
DR EMBL: M84599; AAD40482.1; JOINED.
DR EMBL: M84588; AAD40482.1; JOINED.
DR EMBL: M84589; AAD40482.1; JOINED.
DR EMBL: M84590; AAD40482.1; JOINED.
DR EMBL: M84591; AAD40482.1; JOINED.
DR EMBL: BC000485; AAH00485.1; -
DR EMBL: BC008366; AAH08366.1; -
DR EMBL: L05075; AAB59432.1; -
DR EMBL: S46516; AAB23675.1; -
DR PIR: A33663; DCHUA.
DR MIM: 107930; -
DR InterPro: IPR002129; Pyridoxal dec.
DR Pfam: PF00282; pyridoxal dec. 1.
DR PRINTS: PR00800; YHDCRBOXYLASE.
DR PROSITE: PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase: Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate;
Repeat.
FT BINDING 303 303 PYRIDOXAL PHOSPHATE.
FT DOMAIN 58 178 2 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 58 115 1.
FT REPEAT 118 178 2.
FT CONFLICT 49 49 E -> G (IN REF. 5).
FT CONFLICT 155 155 A -> P (IN REF. 3).
SQ SEQUENCE 480 AA; 53894 MW; 898FD4B9A96C13E CRC64;

Query Match 100.0%; Score 31; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
Db 140 GGGVIO 145

RESULT 6
DDC_MOUSE STANDARD; PRT; 480 AA.
AC O88533;
DT 30-MAY-2000 (Rel. 39, Created)

```

```

DT 30-MAR-2000 (Rel. 39, last sequence update)
DE 01-MAR-2002 (Rel. 41, last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
DE decarboxylase) (DDC).
GN DDC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS WEBSTER;
RA Ducommun A.M., Quach T.T., Gudenithlu K.F., Hadjiconstantinou M.,
RA Neff N.H.;
RT "Cloning of mouse brain aromatic-L-amino-acid decarboxylase cDNA.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLMALANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: CATECHOLAMINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF071068; AAC25566.1; -.
DR MGJ; GGI:94876; Ddc.
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; Pyridoxal_dec.1.
DR PRINTS; PR00800; YHDCRBOXLASE.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase; Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate;
KW Repeat.
FT BINDING 303 303 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT DOMAIN 58 178 2 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 58 115 1.
FT REPEAT 118 178 2.
SQ SEQUENCE 480 AA; 53874 MW; D6C1CA504AC2D10E CRC64;

```

```

Query Match 100.0%; Score 31; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GGGVIO 6
Db 140 GGGVIO 145

```

```

RESULT 7
DDC_RAT STANDARD; PRT; 480 AA.
ID DDC_RAT
AC P14173;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, last sequence update)
DT 01-MAR-2002 (Rel. 41, last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
DE decarboxylase) (DDC).
GN DDC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;

```

```

RX MEDLINE=90046754; PubMed=2813383;
RA Tanaka T., Horio Y., Taketoshi M., Iinamura I., Ando-Yamamoto M.,
RA Kangawa K., Matsuo H., Kuroda M., Wada H.;
RT "Molecular cloning and sequencing of a cDNA of rat dopa
RT decarboxylase: partial amino acid homologues with other enzymes
RT synthesizing catecholamines."
RL Proc. Natl. Acad. Sci. U.S.A. 86:8142-8146(1989).
RN [2]
RP MUTAGENESIS.
RX MEDLINE=97044752; PubMed=8889823;
RA Ishii S., Mizuguchi H., Nishino J., Hayashi H., Kagamiyama H.;
RT "Functionally important residues of aromatic L-amino acid
RT decarboxylase probed by sequence alignment and site-directed
RT mutagenesis."
RL J. Biochem. 120:369-376(1996).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLMALANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: CATECHOLAMINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J33001; AAA40646.1; -.
DR EMBL; J32989; AAA40646.1; JOINED.
DR EMBL; J32990; AAA40646.1; JOINED.
DR EMBL; J32991; AAA40646.1; JOINED.
DR EMBL; J32992; AAA40646.1; JOINED.
DR EMBL; J32993; AAA40646.1; JOINED.
DR EMBL; J32994; AAA40646.1; JOINED.
DR EMBL; J32995; AAA40646.1; JOINED.
DR EMBL; J32996; AAA40646.1; JOINED.
DR EMBL; J32997; AAA40646.1; JOINED.
DR EMBL; J33003; AAA40646.1; JOINED.
DR EMBL; J32999; AAA40646.1; JOINED.
DR EMBL; J33000; AAA40646.1; JOINED.
DR EMBL; M27716; AAA41087.1; -.
DR PIR; A33994; DCRTA.
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; Pyridoxal_dec.1.
DR PRINTS; PR00800; YHDCRBOXLASE.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase; Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate;
KW Repeat.
FT BINDING 303 303 PYRIDOXAL PHOSPHATE.
FT DOMAIN 58 178 2 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 58 115 1.
FT REPEAT 118 178 2.
SQ SEQUENCE 480 AA; 54053 MW; 1E1D077488704574 CRC64;

```

```

Query Match 100.0%; Score 31; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GGGVIO 6
Db 140 GGGVIO 145

```

```

RESULT 8
DDC_PIG STANDARD; PRT; 486 AA.
ID DDC_PIG
AC P80041;
DT 01-NOV-1991 (Rel. 20, Created)

```

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
GN decarboxylase) (DDC).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96207588; PubMed=8670114;
RA Moore P.S., Dominici P., Borri-Voltattorni C.;
RT Cloning and expression of pig kidney dopa decarboxylase: comparison
RT of the naturally occurring and recombinant enzymes.";
RL Biochem. J. 315:249-256(1996).
RN [2]
RP SEQUENCE OF 1-485.
RC TISSUE=Kidney;
RX MEDLINE=92037590; PubMed=1935935;
RA Maras B., Dominici P., Barra D., Bossa F., Borri-Voltattorni C.;
RT Pig kidney 3,4-dihydroxyphenylalanine (dopa) decarboxylase. Primary
RT structure and relationships to other amino acid decarboxylases.";
RL Eur. J. Biochem. 201:385-391(1991).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLALANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: CATECHOLAMINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S82290; AABA7157.1; -.
DR PIR: S17848; S17848.
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec: 1.
DR PRINTS: PR00800; YHDCRBOXLASE.
DR PROSITE: PS00392; DDC_GAD_HDC_YDC: 1.
KM Lyase: Decarboxylase: Catecholamine biosynthesis; Pyridoxal phosphate;
KM Repeat: Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT BINDING 303 303 PYRIDOXAL PHOSPHATE.
FT DOMAIN 58 178 2 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 58 115 1.
FT REPEAT 118 178 2.
SO SEQUENCE 486 AA; 53936 MW; 6CE5978531A9FFA4 CRC64;

```

```

Query Match 100.0%; Score 31; DB 1; Length 486;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGVIO 6
DB 140 GGGVIO 145

```

```

RESULT 9
DDC_BOVIN STANDARD; PRT; 487 AA.
AC P27718;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA

```

```

DE decarboxylase) (DDC).
GN DDC.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90339966; PubMed=2166204;
RA Kang U., Joh T.H.;
RT "Deduced amino acid sequence of bovine aromatic L-amino acid
RT decarboxylase: homology to other decarboxylases.";
RL Brain Res. Mol. Brain Res. 8:83-87(1990).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLALANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: CATECHOLAMINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M74029; AAC41615.1; -.
DR PIR: A43758; A43758.
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec: 1.
DR PRINTS: PR00800; YHDCRBOXLASE.
DR PROSITE: PS00392; DDC_GAD_HDC_YDC: 1.
KM Lyase: Decarboxylase: Catecholamine biosynthesis; Pyridoxal phosphate;
KM Repeat: Acetylation.
FT MOD_RES 303 303 PYRIDOXAL PHOSPHATE.
FT DOMAIN 58 178 2 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 58 115 1.
FT REPEAT 118 178 2.
SO SEQUENCE 487 AA; 54434 MW; EDCE1A7F37DD9D1A CRC64;

```

```

Query Match 100.0%; Score 31; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGVIO 6
DB 140 GGGVIO 145

```

```

RESULT 10
DDC_CATRO STANDARD; PRT; 500 AA.
AC P17770;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
DE decarboxylase) (Tryptophan decarboxylase).
GN TDC.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; easterids I; Gentianales; Apocynaceae; Rauvolfiaceae;
OC Vincae; Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202373; PubMed=2704736;

```

```

RA de Luca V., Marineau C., Brisson N.;
RT "Molecular cloning and analysis of cDNA encoding a plant tryptophan
RT decarboxylase: comparison with animal dopa decarboxylases.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:2582-2586(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, MORNING MIST; TISSUE=Leaf;
RX MEDLINE=94211212; PubMed=8159173;
RA Goddijn O.J.M., Lohman F.P., de Kam R.J., Schlipperoot R.A.,
RA Hoge J.H.C.;
RT "Nucleotide sequence of the tryptophan decarboxylase gene of
RT Catharanthus roseus and expression of tdc-gusa gene fusions in
RT Nicotiana tabacum.";
RL Mol. Genet. 242:217-225(1994).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLAANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25151; AAA33109.1; -
DR EMBL: X67662; CAA47896.1; -
DR EMBL: A23287; CAA01667.1; -
DR PIR: A32103; DCJAP.
DR PIR: S25471; S25471.
DR PIR: S41726; S41726.
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec. 1.
DR PRINTS: PR00800; YHDCBOXLASE.
DR PROSITE: PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase; Decarboxylase; Pyridoxal phosphate.
FT BINDING 319 319 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 500 AA; 56206 MW; 32965957DEC566E7 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
Db 159 GGGVIO 164

RESULT 11
ID DDC_MANSE STANDARD; PRT; 508 AA.
AC P48861;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
DE decarboxylase) (DDC).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingodea; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Epidermis;
RX MEDLINE=95269874; PubMed=7750638;
RA Hiruma K., Carter M.S., Riddiford L.M.;
RA "Characterization of the dopa decarboxylase gene of Manduca sexta and

```

```

RT its suppression by 20-hydroxyecdysone.";
RL Dev. Biol. 169:195-209(1995).
CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN = TRYPTAMINE + CO(2) (ALSO ACTS
CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLAANINE (DOPA)).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U03909; AAC46604.1; -
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec. 1.
DR PRINTS: PR00800; YHDCBOXLASE.
DR PROSITE: PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase; Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate.
FT BINDING 303 303 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 508 AA; 57891 MW; 8753894956A7C77 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
Db 140 GGGVIO 145

RESULT 12
ID DDC_DROME STANDARD; PRT; 510 AA.
AC P05031; P05032; Q24295; O18379; Q9V1Z5; Q9V1Z6;
DT 13-AUG-1987 (Rel. 05, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aromatic-L-amino-acid decarboxylase (EC 4.1.1.28) (AADC) (DOPA
DE decarboxylase) (DDC).
GN DDC OR CG10697.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND ALTERNATIVE SPLICING.
RC STRAIN=CANTON-S;
RX MEDLINE=87133496; PubMed=3102230;
RA Morgan B.A., Johnson W.A., Hirsh J.;
RT "Regulated splicing produces different forms of dopa decarboxylase in
RT the central nervous system and hypoderm of Drosophila melanogaster.";
RL EMBO J. 5:3335-3342(1986).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=CANTON-S;
RX MEDLINE=87053836; PubMed=3023054;
RA Eysleth D.D., Gietz R.D., Spencer C.A., Nargang F.E., Hodgetts R.B.,
RA Marsh J.L.;
RT "Sequence and structure of the dopa decarboxylase gene of Drosophila:
RT evidence for novel RNA splicing variants.";
RL EMBO J. 5:2663-2672(1986).
RN [3]
RP REVISIONS.
RC STRAIN=CANTON-S;
RA Marsh J.L.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

```

[4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Goebye J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins K.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champs M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikiel G.L.G.,
 RA April J.F., Adganyi A., An H.-J., Andrews-Ffankoch C., Baldwin D.,
 RA Ballou R.M., Besu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Meriklov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacelb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sledge-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN
 RP SEQUENCE OF 74-510 FROM N.A.
 RC STRAIN-ST. LUCIA;
 RX MEDLINE-99250256; PubMed-9950256;
 RA Tatarenkov A., Saez A.G., Ayala F.J.;
 RT "A compact gene cluster in *Drosophila*: the unrelated Cs gene is
 RT compressed between duplicated amd and Ddc.";
 RL Gene 231:111-120(1999).
 RN
 RP SEQUENCE OF 493-510 FROM N.A.
 RX MEDLINE-88038375; PubMed-3478553;
 RA Eveleth D.D., Marsh J.L.;
 RT "Overlapping transcription units in *Drosophila*: Sequence and structure
 RT of the Cs gene.";
 RL Mol. Gen. Genet. 209:290-298(1987).
 RN
 RP TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RX MEDLINE-87042790; PubMed-3095924;
 RA Scholnick S.B., Bray S.J., Morgan B.A., McCormick C.A., Hirsch J.;
 RT "CNS and hypoderm regulatory elements of the *Drosophila melanogaster*
 RT dopa decarboxylase gene.";
 RL Science 234:998-1002(1986).
 RN
 RP DEVELOPMENTAL STAGE.
 RX MEDLINE-86165362; PubMed-3007242;
 RA Spencer C.A., Gietz R.D., Hodgetts R.B.;
 RT "Analysis of the transcription unit adjacent to the 3'-end of the dopa
 RT decarboxylase gene in *Drosophila melanogaster*.";
 RL Dev. Biol. 114:260-264(1986).
 RN
 RP INDUCTION.

RX MEDLINE-86165352; PubMed-3007239;
 RA Clark W.C., Doctor J., Fristrom J.W., Hodgetts R.B.;
 RT "Differential responses of the dopa decarboxylase gene to
 RT 20-OH-ecdysone in *Drosophila melanogaster*.";
 RL Dev. Biol. 114:141-150(1986)
 CC
 CC -1- CATALYTIC ACTIVITY: L-TRYPTOPHAN + TRYPHTAMINE + CO(2) (ALSO ACTS
 CC ON 5-HYDROXY-L-TRYPTOPHAN AND DIHYDROXY-L-PHENYLALANINE (DOPA)).
 CC
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC
 CC -1- SUBUNIT: HOMODIMER.
 CC
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CNS/LONG (SHOWN HERE) AND
 CC HYPODERM/SHORT: ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC
 CC -1- TISSUE SPECIFICITY: HYPODERM ISOFORM IS EXPRESSED ONLY IN
 CC HYPODERMAL EPITHELIUM AND THE CNS ISOFORM ONLY IN CENTRAL NERVOUS
 CC SYSTEM.
 CC
 CC -1- DEVELOPMENTAL STAGE: HIGH EXPRESSION LEVELS IN HYPODERM DURING
 CC LATE EMBRYOGENESIS, LATE LARVAL DEVELOPMENT, PUPARIATION AND ADULT
 CC ECLOSION. CONSTANT EXPRESSION LEVEL IN CNS THROUGHOUT THE LIFE
 CC CYCLE.
 CC
 CC -1- INDUCTION: BY ECDYSONE. IN LARVAL EPIDERMIS, EXPRESSION IS RAPIDLY
 CC INDUCED. IN ADULT EPIDERMIS EXPRESSION RESPONDS TO A PULSE OF
 CC HORMONE AND THERE IS A TIME LAG BETWEEN INITIAL EXPOSURE AND
 CC APPEARANCE OF DDC.
 CC
 CC -1- SIMILARITY: BELONGS TO GROUP II: DECARBOXYLASES (DDC, GAD, HDC AND
 CC TYRDC).
 CC
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: X04661; CAB37087.1; -;
 DR EMBL: X04661; CAB37088.1; -;
 DR EMBL: X04426; CAA28022.1; -;
 DR EMBL: X04426; CAA28023.1; -;
 DR EMBL: AE003661; AAF53762.1; -;
 DR EMBL: AE003661; AAF53763.1; -;
 DR EMBL: AF091328; AAC67582.1; -;
 DR EMBL: X05991; CAA29409.2; -;
 DR PIR: A25709; DCEFA.
 DR PIR: A25697; DCEFD1.
 DR PIR: B25697; DCEFD2.
 DR Flybase: FBgn0000422; Ddc.
 DR InterPro: IPR002129; Pyridoxal dec.
 DR Pfam: PF00282; Pyridoxal dec. 1.
 DR PRINTS: PR00800; TYRDCROXLASe.
 DR PROSITE: PS00392; DDC_GAD_HDC_YDC; 1.
 DR Lyase: Decarboxylase; Catecholamine biosynthesis; Pyridoxal phosphate;
 KW Alternative splicing.
 FT BINDING 337 337 PYRIDOXAL PHOSPHATE.
 FT BINDING 337 35 MISSING (IN HYPODERM ISOFORM).
 FT BINDING 1 33 MSHPISNTIPTKOTDNGKANSIPKLDPKYS -> MSIG
 FT BINDING 1 33 PRIRANNYARLITKFCILIK (IN REF. 2;
 FT CONFLICT 32 33 CAA28023).
 FT CONFLICT 479 479 MISSING (IN REF. 1).
 FT CONFLICT 479 479 R -> A (IN REF. 1).
 SQ SEQUENCE 510 AA; 57287 MW; 0A850488D407DABF CRC64;
 Query Match 100.0%; Score 31; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GGGVIG 6
 Db 175 GGGVIG 180
 RESULT 13
 LZAM_DROME STANDARD; PRT; 510 AA.
 ID LZAM_DROME

CC P18486: 09V128;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Alpha-methylidopa hypersensitive protein (EC 4.1.1.-).
 GN AND OR I(2)AMD OR CG10501.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Phryganea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP MEDLINE=67031534; PubMed=3021571;
 RA Marsh J.L., Erile M.P., Leeds C.A.;
 RT "Molecular localization, developmental expression and nucleotide
 sequence of the alpha-methylidopa hypersensitive gene of Drosophila.";
 RL Genetics 114:453-467(1986).
 RN [2]
 RP REVISIONS.
 RA Marsh J.L.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Adayant A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moadary C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: PARTICIPATES IN CATECHOLAMINE METABOLISM. IT PLAYS A
 VITAL ROLE IN CUTICLE DEVELOPMENT.
 CC -1- CORFACTOR: PYRIDOXAL PHOSPHATE (Potential).
 CC -1- DEVELOPMENTAL STAGE: REACHES A MAXIMUM IN MID-EMBRYOGENESIS.
 CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 TYRDC).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X04695; CAA28400.1; -;
 DR EMBL: AE003661; AAF53760.1; -;
 DR PIR: A28569; A28569.
 DR PIRbase: FBgn0000075; amd.
 DR InterPro: IPR002129; Pyridoxal.dec.
 DR PRINTS: PF00282; Pyridoxal.dec.1.
 DR PROSITE: PS00392; DDC_GAD_HDC_YDC_1.
 KW Catecholamine metabolism; Lyase; Decarboxylase; Pyridoxal phosphate;
 KW Cuticle.
 FT BINDING 303 303 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT CONFLICT 71 72 WQ -> SE (IN REF. 1).
 FT CONFLICT 370 370 G -> E (IN REF. 1).
 FT CONFLICT 379 379 R -> A (IN REF. 1).
 FT CONFLICT 477 478 EQ -> DE (IN REF. 1).
 FT CONFLICT 493 494 OH -> HD (IN REF. 1).
 SO SEQUENCE 510 AA; 57021 MW; 94464F49896E7654 CRC64;

Query Match 100.0%; Score: 31; DB 1; Length 510;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIG 6
 DB 140 GGGVIG 145

RESUME 14
 HV31_HUMAN STANDARD; PRT; 119 AA.
 AC P01770;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig heavy chain V-III region NIE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=77070269; PubMed=826475;
 RA Ponsling H., Hilschmann N.;
 RT "The rule of antibody structure. The primary structure of a
 RT monoclonal IgG1 immunoglobulin (myeloma protein Nle). III. The
 RT chymotryptic peptides of the H-chain, alignment of the tryptic
 RT peptides and discussion of the complete structure.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
 RN [2]
 RP DISULFIDE BOND.
 RX MEDLINE=77070267; PubMed=1002129;
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 RT IgG1 immunoglobulin (myeloma protein Nle). I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT cyanogen bromide cleavage products, and the disulfide bridges.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IG G1 MYELOMA
 PROTEIN.
 CC PIR: A02053; GIHMNT.
 DR HSSP: P01772; 2IG2.
 DR InterPro: IPR003006; Ig MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IgV; 1.
 KW Immunoglobulin V region.

FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 22 96
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA; 13242 MW; C96935A6E5E165B CRC64;

Query Match

Best Local Similarity 96.8%; Score 30; DB 1; Length 119;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
 1111:1
 Db 8 GGGVIO 13

RESULT 15

HV3L_HUMAN
 ID HV3L_HUMAN STANDARD; PRT; 119 AA.
 AC P01773;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, last sequence update)
 DT 15-JUL-1999 (Rel. 38, last annotation update)
 DE Ig heavy chain V-III region BUR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE (MYELOMA PROTEIN BUR).
 RX MEDLINE=79151016; PubMed=107164;
 RA Putnam F.W., Liu Y.-S.V., Low T.L.K.;
 RT "Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal
 RT IgA1 protease, digestion, Fab and Fc fragments, and the complete
 RT amino acid sequence of the alpha 1 heavy chain.";
 RL J Biol Chem. 254:2865-2874(1979).
 DR PIR; A02056; A1HDB.
 DR HSSP; P01772; 2IG2.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00406; IGV; 1.
 KW Immunoglobulin V region; Glycoprotein.
 FT MOD_RES 1 1
 FT DISULFID 22 96
 FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .).
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA; 12981 MW; 12A709A75344D024 CRC64;

Query Match

Best Local Similarity 96.8%; Score 30; DB 1; Length 119;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIO 6
 1111:1
 Db 8 GGGVIO 13

Search completed: May 24, 2002, 09:57:41
 Job time: 330 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:51:41 ; Search time 41.33 Seconds
(without alignments)
25.114 Million cell updates/sec

Title: US-09-913-763-1
Perfect score: 31
Sequence: 1 GGGVIG 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	141	5	Q9BNZ3
2	31	100.0	173	5	Q9NWX3
3	31	100.0	173	5	Q9NWX2
4	31	100.0	173	5	Q9NWX7
5	31	100.0	183	5	Q9NWX1
6	31	100.0	183	5	Q9NWX9
7	31	100.0	189	5	Q9NWX5
8	31	100.0	202	5	Q9NWX4
9	31	100.0	202	5	Q9NWX0
10	31	100.0	217	5	Q9Y1T0
11	31	100.0	217	5	Q9Y1R7
12	31	100.0	217	5	Q9Y0Q7
13	31	100.0	226	5	Q9NWX8
14	31	100.0	226	5	Q9NWX6
15	31	100.0	226	5	Q9NWX5
16	31	100.0	226	5	Q9NWX4

17	31	100.0	226	5	Q9Y0S6	Q9Y0S6 datana pers
18	31	100.0	232	5	Q9Y0Q8	Q9Y0Q8 meganola sp
19	31	100.0	234	5	Q9NWX6	Q9NWX6 cardioladi
20	31	100.0	235	5	Q9Y0S3	Q9Y0S3 nerice bide
21	31	100.0	236	5	Q9Y1S9	Q9Y1S9
22	31	100.0	236	5	Q9Y1S8	Q9Y1S8 adela purpu
23	31	100.0	236	5	Q9Y1S7	Q9Y1S7 brachycentr
24	31	100.0	236	5	Q9Y1S6	Q9Y1S6 cecidoses e
25	31	100.0	236	5	Q9Y1S5	Q9Y1S5 copridosca
26	31	100.0	236	5	Q9Y1S4	Q9Y1S4 chalcocopa
27	31	100.0	236	5	Q9Y1S2	Q9Y1S2 epimartyrila
28	31	100.0	236	5	Q9Y1S1	Q9Y1S1 eriocranila
29	31	100.0	236	5	Q9Y1S0	Q9Y1S0 greya obscu
30	31	100.0	236	5	Q9Y1R9	Q9Y1R9 greya pollt
31	31	100.0	236	5	Q9Y1R5	Q9Y1R5 olieira sp.
32	31	100.0	236	5	Q9Y1R4	Q9Y1R4 palaeophatus
33	31	100.0	236	5	Q9Y1R1	Q9Y1R1 dolophillode
34	31	100.0	236	5	Q9Y1R0	Q9Y1R0 sabatinnca z
35	31	100.0	236	5	Q9Y1Q9	Q9Y1Q9 tischeria b
36	31	100.0	236	5	Q9Y1Q7	Q9Y1Q7 tischeria c
37	31	100.0	236	5	Q9Y1Q6	Q9Y1Q6 thyriddopter
38	31	100.0	236	5	Q9Y1Q4	Q9Y1Q4 tetragma ge
39	31	100.0	236	5	Q9Y1Q3	Q9Y1Q3 linea pellil
40	31	100.0	236	5	Q9Y1Q1	Q9Y1Q1 vespina que
41	31	100.0	236	5	Q9Y0S7	Q9Y0S7 furcula cin
42	31	100.0	236	5	Q9Y0N0	Q9Y0N0 catabena ii
43	31	100.0	236	5	Q9BNZ7	Q9BNZ7 aellopos ta
44	31	100.0	236	5	Q9BNZ1	Q9BNZ1 eumorpia pa
45	31	100.0	236	5	Q9BNY9	Q9BNY9 hemaris thy

ALIGNMENTS

RESULT 1
Q9BNZ3 PRELIMINARY; PRT; 141 AA.
ID Q9BNZ3
AC Q9BNZ3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DOPA DECARBOXYLASE (FRAGMENT).
OS Dryocampa rubicunda (rosy maple moth).
OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
OC Pterygota: Neoptera: Endopterygota: Lepidoptera: Glossata: Ditrysia;
OC Bombycoidea: Saturniidae; Citheroniinae; Dryocampa.
OX NCBI_TaxID=119254;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21370235; PubMed=11476639;
RX Reigler J.C., Mitter C., Friedlander T.P., Reigler R.S.;
RT "Re: phylogenetic relationships in Springidae (Insecta: Lepidoptera):
RT Initial evidence from two nuclear genes."
RL Mol. Phylogenet. Evol. 20:311-316(2001).
CC -i- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -i- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL: AF234586; AA08683.1; -;
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; pyridoxal_dec.1.
KW Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1
FT NON_TER 141
SQ SEQUENCE 141 AA; 15013 MW; 8A38056F802BE511 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 141;

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIG 6
Db 15 GGGVIG 20

```

RESULT 2
Q9NUX3
ID Q9NUX3 PRELIMINARY; PRT; 173 AA.
AC Q9NUX3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DOPA DECARBOXYLASE (FRAGMENT).
GN DDC.
OS Helodon onychodactylus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Simuliidae; Helodon.
OX NCBI_TaxID=62327;
RN [1]
RP SEQUENCE FROM N.A.
RT Moulton J.K.;
RL "Molecular sequence data resolves basal divergences within Simuliidae (Diptera).";
RL Syst. Entom. 25:95-113(2000).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL; AF078879; AAF30392.1; -
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec; 1.
DR PRINTS; PR00800; YHDCRBOXLASE.
KM Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1
FT NON_TER 173
SQ SEQUENCE 173 AA; 18183 MW; 36FAF3237B401999 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIQ 6
Db 37 GGGVIQ 42

RESULT 3
Q9NUX2
ID Q9NUX2 PRELIMINARY; PRT; 173 AA.
AC Q9NUX2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DOPA DECARBOXYLASE (FRAGMENT).
GN DDC.
OS Prosimulium formosum.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Simuliidae; Prosimulium.
OX NCBI_TaxID=62326;
RN [1]
RP SEQUENCE FROM N.A.
RT Moulton J.K.;
RL "Molecular sequence data resolves basal divergences within Simuliidae (Diptera).";
RL Syst. Entom. 25:95-113(2000).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL; AF078880; AAF30393.1; -
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec; 1.
DR PRINTS; PR00800; YHDCRBOXLASE.
KM Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1
FT NON_TER 173
SQ SEQUENCE 173 AA; 18349 MW; 7FBE784166116BDC CRC64;

```

```

FT NON_TER 173
SQ SEQUENCE 173 AA; 18329 MW; 37219E85A568F9E7 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIQ 6
Db 37 GGGVIQ 42

RESULT 4
Q9NUJ7
ID Q9NUJ7 PRELIMINARY; PRT; 173 AA.
AC Q9NUJ7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DOPA DECARBOXYLASE (FRAGMENT).
GN DDC.
OS Crozetia crozetensis.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Simuliidae; Crozetia.
OX NCBI_TaxID=61033;
RN [1]
RP SEQUENCE FROM N.A.
RT Moulton J.K.;
RL "Molecular sequence data resolves basal divergences within Simuliidae (Diptera).";
RL Syst. Entom. 25:95-113(2000).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL; AF078885; AAF30398.1; -
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec; 1.
DR PRINTS; PR00800; YHDCRBOXLASE.
KM Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1
FT NON_TER 173
SQ SEQUENCE 173 AA; 18349 MW; 7FBE784166116BDC CRC64;

Query Match 100.0%; Score 31; DB 5; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVIQ 6
Db 37 GGGVIQ 42

RESULT 5
Q9NUX1
ID Q9NUX1 PRELIMINARY; PRT; 183 AA.
AC Q9NUX1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DOPA DECARBOXYLASE (FRAGMENT).
GN DDC.
OS Prosimulium impostor.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomidae; Simuliidae; Prosimulium.
OX NCBI_TaxID=87546;
RN [1]
RP SEQUENCE FROM N.A.
RT Moulton J.K.;
RL "Molecular sequence data resolves basal divergences within Simuliidae (Diptera).";
RL Syst. Entom. 25:95-113(2000).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL; AF078885; AAF30398.1; -
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec; 1.
DR PRINTS; PR00800; YHDCRBOXLASE.
KM Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1
FT NON_TER 173
SQ SEQUENCE 173 AA; 18349 MW; 7FBE784166116BDC CRC64;

```

RT (Diptera).";
RL Syst. Entom. 25:95-113(2000).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL: AF078881; AAF30394.1; -
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec; 1.
DR PRINTS: PR00800; YHDCRBOXYLASE.
KM Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1 183
SQ SEQUENCE 183 AA; 19461 MW; 2B6659504072E741 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
Db 37 GGGVITQ 42

RESULT 6
Q9NJW9 PRELIMINARY; PRT; 183 AA.
AC Q9NJW9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE DOPA DECARBOXYLASE (FRAGMENT).
GN DDC.
OS Urosimulium aculeatum.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomoidae; Simuliidae; Urosimulium.
OX NCBI_TaxID=61073;
RN [1]
RP SEQUENCE FROM N.A.
RA Moulton J.K.;
RT "Molecular sequence data resolves basal divergences within Simuliidae (Diptera).";
RL Syst. Entom. 25:95-113(2000).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL: AF078883; AAF30396.1; -
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec; 1.
DR PRINTS: PR00800; YHDCRBOXYLASE.
KM Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1 183
SQ SEQUENCE 183 AA; 19479 MW; C5B79161CA44A4C9 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
Db 37 GGGVITQ 42

RESULT 7
Q9NXX5 PRELIMINARY; PRT; 189 AA.
AC Q9NXX5;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
OY 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE DOPA DECARBOXYLASE (FRAGMENT).
GN DDC.
OS Parasimulium crosskeyi.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomoidae; Simuliidae; Parasimulium.
OX NCBI_TaxID=61060;
RN [1]
RP SEQUENCE FROM N.A.
RA Moulton J.K.;
RT "Molecular sequence data resolves basal divergences within Simuliidae (Diptera).";
RL Syst. Entom. 25:95-113(2000).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL: AF078877; AAF30390.1; -
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec; 1.
DR PRINTS: PR00800; YHDCRBOXYLASE.
KM Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1 189
SQ SEQUENCE 189 AA; 20217 MW; 40EBBBB59488904 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6
Db 43 GGGVITQ 48

RESULT 8
Q9NXX4 PRELIMINARY; PRT; 202 AA.
AC Q9NXX4;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE DOPA DECARBOXYLASE (FRAGMENT).
GN DDC.
OS Gymnopsis fimbriatus.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomoidae; Simuliidae; Gymnopsis.
OX NCBI_TaxID=61043;
RN [1]
RP SEQUENCE FROM N.A.
RA Moulton J.K.;
RT "Molecular sequence data resolves basal divergences within Simuliidae (Diptera).";
RL Syst. Entom. 25:95-113(2000).
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).
DR EMBL: AF078878; AAF30391.1; -
DR InterPro: IPR002129; Pyridoxal_dec.
DR Pfam: PF00282; Pyridoxal_dec; 1.
DR PRINTS: PR00800; YHDCRBOXYLASE.
KM Decarboxylase; Lyase; Pyridoxal phosphate.
FT NON_TER 1 202
SQ SEQUENCE 202 AA; 21828 MW; F30E49A8744A59AF CRC64;

Query Match 100.0%; Score 31; DB 5; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVITQ 6

Db 37 GGGVIO 42

RESULT 9

09NXX0 PRELIMINARY; PRT; 202 AA.

ID 09NXX0;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE DOPA DECARBOXYLASE (FRAGMENT).

GN DDC.

OS *Twinnia nova*.

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;

OC Chironomidae; Simuliidae; *Twinnia*.

OX NCBI_TaxID=61071;

RP SEQUENCE FROM N.A.

RA Moulton J.K.;

RT "Molecular sequence data resolves basal divergences within Simuliidae (Diptera).";

RL Syst. Entom. 25:95-113(2000).

CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).

DR EMBL; AF078882; AAF30395.1; -

DR InterPro; IPR002129; Pyridoxal_dec.

DR Pfam; PF00282; Pyridoxal_dec.1.

DR PRINTS; PR00800; YHDCRBOXLASE.

KW Decarboxylase; Lyase; Pyridoxal phosphate.

FT NON_TER 1

FT NON_TER 202

SO SEQUENCE 202 AA; 21377 MW; E00B3D80331EA678 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 202;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
Db 45 GGGVIO 50

RESULT 10

09Y1T0 PRELIMINARY; PRT; 217 AA.

ID 09Y1T0;

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE DOPA DECARBOXYLASE (FRAGMENT).

OS *Antispila cornifolia*.

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;

OC Platyphoridae; Heliozelidae; *Antispila*.

OX NCBI_TaxID=98955;

RP SEQUENCE FROM N.A.

RA Friedlander T.P., Regier J.C., Mitter C., Wagner D.L.;

RT "Relationships within Heteroneuran Lepidoptera: Evidence from Dopa Decarboxylase.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE GNTN FAMILY OF TRANSCRIPTIONAL REGULATORS.

CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).

DR EMBL; AF093166; AAD42709.1; -

DR InterPro; IPR000524; HTH_GntR.

DR InterPro; IPR002129; Pyridoxal_dec.

DR Pfam; PF00282; Pyridoxal_dec.1.
DR PRINTS; PR00035; HTHGNTN.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC.1.
KW DNA-binding; Decarboxylase; Lyase; Pyridoxal phosphate;
KW Transcription regulation.
FT NON_TER 1
FT NON_TER 217
SO SEQUENCE 217 AA; 23788 MW; 63946A5E1C0E6D32 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
Db 15 GGGVIO 20

RESULT 11

09Y1R7 PRELIMINARY; PRT; 217 AA.

ID 09Y1R7;

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE DOPA DECARBOXYLASE (FRAGMENT).

OS *Lampronia aeneus*.

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;

OC Incurvatiidae; Prodoxidae; *Lampronia*.

OX NCBI_TaxID=30234;

RP SEQUENCE FROM N.A.

RA Friedlander T.P., Regier J.C., Mitter C., Wagner D.L.;

RT "Relationships within Heteroneuran Lepidoptera: Evidence from Dopa Decarboxylase.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE GNTN FAMILY OF TRANSCRIPTIONAL REGULATORS.

CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND TYRDC).

DR EMBL; AF093179; AAD42722.1; -

DR InterPro; IPR000524; HTH_GntR.

DR Pfam; PF00282; Pyridoxal_dec.1.

DR PRINTS; PR00035; HTHGNTN.

DR PROSITE; PS00392; DDC_GAD_HDC_YDC.1.

KW DNA-binding; Decarboxylase; Lyase; Pyridoxal phosphate;

KW Transcription regulation.

FT NON_TER 1

FT NON_TER 217

SO SEQUENCE 217 AA; 23697 MW; 20853B9B15D4522 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
Db 15 GGGVIO 20

RESULT 12

09Y007 PRELIMINARY; PRT; 217 AA.

ID 09Y007;

DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE DOPA DECARBOXYLASE (FRAGMENT).

GN DDC.
 OS Baileya levitans.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Noctuidae; Noctuidae; Sarrithripinae; Baileya.
 OX NCBI_TaxID=56374;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mitchell A., Mitter C., Regier J.C.;
 RT "More characters or more taxa revisited: combining data from nuclear
 RT protein-encoding genes for phylogenetic analyses of Noctuoidea
 (Insecta: Lepidoptera).";
 RL Syst. Biol. 49:202-224(2000).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GNTM FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 CC TYRDC).
 DR EMBL: AF151559; AAD37642.1; -.
 DR InterPro: IPR000524; HTM_GntR.
 DR InterPro: IPR002129; Pyridoxal_dec.
 DR Pfam: PF00282; Pyridoxal_dec; 1.
 DR PRINTS: PR00035; HTMGNTM.
 DR PROSITE: PS00392; DDC_GAD_HDC_YDC; 1.
 KM DNA-binding; Decarboxylase; Lyase; Pyridoxal phosphate;
 KW Transcription regulation.
 FT NON_TER 1 217
 FT NON_TER 1 217
 SO SEQUENCE 217 AA; 23801 MW; C72FDA63550C18A6 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 217;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
 Db 15 GGGVIO 20

RESULT 13
 O9NMJ8 PRELIMINARY; PRT; 226 AA.
 AC O9NMJ8;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE DOPA DECARBOXYLASE (FRAGMENT).
 GN DDC.
 OS Austrosimulium bancrofti.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Simuliidae; Austrosimulium.
 OX NCBI_TaxID=27463;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Moulton J.K.;
 RT "Molecular sequence data resolves basal divergences within Simuliidae
 (Diptera).";
 RL Syst. Entom. 25:95-113(2000).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 CC TYRDC).
 DR EMBL: AF078884; AAF30397.1; -.
 DR InterPro: IPR002129; Pyridoxal_dec.
 DR Pfam: PF00282; Pyridoxal_dec; 1.
 DR PRINTS: PR00800; YHDCRBOXLASE.
 KW Decarboxylase; Lyase; Pyridoxal phosphate.
 FT NON_TER 1 226
 FT NON_TER 1 226
 SO SEQUENCE 226 AA; 24052 MW; A9A93AAA0CA24E5D CRC64;

Query Match 100.0%; Score 31; DB 5; Length 226;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
 Db 69 GGGVIO 74

RESULT 14
 O9NMJ6 PRELIMINARY; PRT; 226 AA.
 AC O9NMJ6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE DOPA DECARBOXYLASE (FRAGMENT).
 GN DDC.
 OS Greniera fabrt.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Simuliidae; Greniera.
 OX NCBI_TaxID=61041;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Moulton J.K.;
 RT "Molecular sequence data resolves basal divergences within Simuliidae
 (Diptera).";
 RL Syst. Entom. 25:95-113(2000).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
 CC TYRDC).
 DR EMBL: AF078886; AAF30399.1; -.
 DR InterPro: IPR002129; Pyridoxal_dec.
 DR Pfam: PF00282; Pyridoxal_dec; 1.
 DR PRINTS: PR00800; YHDCRBOXLASE.
 KW Decarboxylase; Lyase; Pyridoxal phosphate.
 FT NON_TER 1 226
 FT NON_TER 1 226
 SO SEQUENCE 226 AA; 23996 MW; A8192100B8024E5D CRC64;

Query Match 100.0%; Score 31; DB 5; Length 226;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVIO 6
 Db 69 GGGVIO 74

RESULT 15
 O9NMJ5 PRELIMINARY; PRT; 226 AA.
 AC O9NMJ5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE DOPA DECARBOXYLASE (FRAGMENT).
 GN DDC.
 OS Paracnephia thornei.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Simuliidae; Paracnephia.
 OX NCBI_TaxID=61056;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Moulton J.K.;
 RT "Molecular sequence data resolves basal divergences within Simuliidae
 (Diptera).";
 RL Syst. Entom. 25:95-113(2000).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND

CC TYRDC)
 DR EMBL; AF078887; AAF30400.1; -;
 DR InterPro; IPR002129; Pyridoxal_dec.
 DR Pfam; PF00282; Pyridoxal_dec.1.
 DR PRINTS; PR00800; YHDCRBOXLASE.
 KW Decarboxylase; Lyase; Pyridoxal phosphate.
 FT NON_TER 1
 FT NON_TER 226
 SQ SEQUENCE 226 AA; 24028 MW; 847CF80D74AA2EA3 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 226;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVYQ 6
 |||||
 Db 69 GGGVYQ 74

Search completed: May 24, 2002, 09:57:20
 Job time: 339 sec